Front. Biol. 2011, 6(1): 31–39
DOI 10.1007/s11515-011-1090-1

REVIEW

Techniques of cell type-specific transcriptome analysis and applications in researches of sexual plant reproduction

Tian-Xiang HU, Miao YU, Jie ZHAO (✉)

Key Laboratory of the Ministry of Education for Plant Developmental Biology, College of Life Sciences, Wuhan University, Wuhan 430072, China

Abstract **In higher plants, specific cell differentiation and fate decision are controlled by differential gene expression. Cell type-specific transcriptome analysis has become an important tool for investigating cell regulatory mechanisms. In recent years, many different techniques have been developed for the isolation of specific cells and the subsequent transcriptome analysis, and considerable data are available regarding the transcriptional profiles of some specific cells. These cell type-specific transcriptome analyses hold significant promise for elucidating the gene expression linked to cellular identities and functions, and are extraordinarily important for research in functional genomics and systems biology aimed toward basic understanding of molecular networks and pathway interactions. Moreover, to reveal the critical mechanisms about sexual plant reproduction, the gamete and embryo cells have long been treated as good subjects for cell-specific transcriptome analysis, and there has been important progress in recent decades. In this review, we summarize current technologies in cell type-specific transcriptome analysis and review the applications of these technologies in research into the mechanisms of sexual reproduction in higher plants.**

Keywords specific cell, transcriptome, plant reproduction

Introduction

In the multi-cellular organisms of higher plants, different cells are specialized into different identities and so compose the organism. The identities of individual plant cells are specified by response to numerous different components, such as DNA, RNA, proteins and small molecules, which participate in the regulatory networks to coordinate multiple biological functions of cells. A global overview of gene expression will promote the elucidation of the transcriptional networks linked to cellular identity and function. The availability of high-throughput experimental methods has enabled researchers to determine the expression levels for thousands of genes, and promises to provide such cellular information on a genome- and proteome-wide scale. Thus the transcriptome profiles in a particular type of cells can ultimately be analyzed and recorded, and such a fingerprint of gene-expression patterns should provide useful markers for specific cells and ultimately contribute to understanding the molecular mechanism of cell-fate decisions. However, the efficiency of information derived from RNA and protein expression profiling is dependent on the specificity of the biological starting materials. It has been challenging to obtain one kind of cell at a specific developmental stage and/or from a unique location in plants, and thus most transcriptome profiling is carried out with whole-tissue resolution.

Tissue transcriptional profiling analysis has provided important spatial information of gene expression, even though plant tissues almost invariably contain varying proportions of multiple cell types. To investigate the transcriptional profiles of specific cell types, different techniques have been adopted. These techniques are applied respectively at the two different steps: the isolation methods of specific types of cells, and the obtaining and analyzing of cell transcriptomes. Techniques for isolating specific cells mainly include manual and enzymatic micromanipulation, laser microdissection (LMD) and fluorescence activated cell sorting (FACS) (Hu and Yang, 2002; Birnbaum et al., 2003; Nelson et al., 2006). cDNA library construction and DNA microarray are widely used to analyze cell transcriptomes, and the RNA-Seq (high-throughput cDNA sequencing) technique by next-generation (NG) sequencing is another emerging powerful tool (Dresselhaus et al., 1994; Honys and

Received November 3, 2010; accepted December 6, 2010
Correspondence: Jie ZHAO
E-mail: jzhao@whu.edu.cn

Twell, 2004; Tang et al., 2009). In this review, we summarize the current techniques for analyzing the transcriptomes of specific types of cells, especially in male and female gametes, zygotes and early proembryos of higher plants.

Sexual reproduction is an important biological event not only for evolution but also for breeding in higher plants. Plant reproduction occurs through the production of gametes from the gametophyte as a haploid generation. In flowering plants, there are highly reduced male and female gametophytes, the pollen grain and embryo sac, respectively, which consist of only a few cells. During double fertilization, one sperm fuses with an egg cell and another sperm fuses with a central cell to produce embryo and endosperm, respectively (Russell, 1992). Embryo development in higher plants is an important phase of the life cycle, in which the main body plan of a plantlet is established. In this article, we summarize the applications of techniques for investigating cell type-specific transcriptomes in sexual plant reproduction, and review recent progress in understanding the mechanisms underlying gamete and gametophytic generation, egg cell fertilization, zygote activation and asymmetrical division, orientated differentiation of apical and basal cells from two-celled proembryos, and polarity formation in proembryo development in higher plants.

Techniques for isolating specific types of cells

Micromanipulation

Micromanipulation, a powerful skill, is used to isolate single cells from complex tissues and organs, especially for egg cells and zygotes. There are three major isolation methods: direct manual microdissection, enzymatic maceration assisted by shaking or grinding, and the combination of enzymatic maceration and microdissection (Fu et al., 1996; Hu and Yang, 2002). Enzyme digestion combined with shaking requires extended incubation in enzymatic solution, and therefore can affect the vitality of living cells and so has been infrequently adopted. The other two methods, direct manual microdissection and enzyme digestion combined with microdissection, are used more universally in many plants. As early as 1985, viable embryo sacs and their protoplasts in *Nicotiana tabacum* were isolated from ovaries using enzymatic degradation and grinding methods (Hu et al., 1985). The isolation methods for embryo sac cells and zygotes are well developed in a wide range of plant species: e.g., *Plumbago zeylanica* (Huang and Russell, 1989), *Zea mays* (Kranz et al., 1991; Kumlehn et al., 1998), *Triticum aestivum* (Holm et al., 1994), *Brassica napus* (Katoh et al., 1997), *Oryza sativa* (Zhao et al., 2000), *Torenia fournieri* (Chen et al., 2008), *Antirrhinum majus* and *Arabidopsis thaliana* (Wang et al., 2010).

Compared with female gametes, male gamete cells are usually exposed and easily isolated. Grinding and osmotic shock methods are widely used to release viable sperm cells from mature pollen grains, and the sperm cells can then be purified using Percoll density gradient centrifugation or manual collection with a micropipette under a microscope (Hu and Yang, 2002). By using these methods, viable sperm cells have been successfully isolated from species such as tobacco (Cao et al., 1996), *P. zeylanica* (Russell 1986; Zhang et al., 1998), *Lilium longiflorum* (Tanaka, 1988) and maize (Kranz et al., 1991). The isolated specific types of cells were used as samples for cDNA library construction and DNA microarray analysis for investigating the cell specific transcriptomes, which greatly broadened knowledge of gene expression in sperm, egg and zygote cells (Dresselhaus et al., 1994; Lê et al., 2005; Ning et al., 2006; Sprunck et al., 2005; Gou et al., 2009). In addition to the above, other methods, such as microcapillaries or micropipettes have been used to directly obtain the contents of leaf epidermal, mesophyll, companion and guard cells (Karrer et al., 1995; Brandt et al., 1999, 2002). Although micromanipulation is a powerful method for isolating specific types of cells, its application is relatively limited since it is time consuming and requires extensive skills.

Laser microdissection

Laser microdissection (LMD) is a technique based on the application of laser technology. It circumvents some limitations of micromanipulation, and provides a rapid and reliable method to procure pure populations of cells from specific regions of tissue sections. Plant tissues containing cells of interest are first fixed with appropriate fixatives, and then embedded to prepare sections for LMD. The dissection process is accomplished by adapting a laser beam device with the microscopy and software system. The laser dissection device is used to incise the specific types of cells from tissue sections, which are immobilized on a slide under direct view of a microscope (Emmert-Buck et al., 1996; Simone et al., 1998). The isolated cell samples can then be collected by various methods, for example, with an adhesive film in conjunction with a collector tube under the slide (Kerk et al., 2003), or by catapulting the cells with laser pulses and harvesting them in a collector above the sample (Nakazono et al., 2003).

The first report of LMD used in plant cells was for isolating phloem cells of rice (Asano et al., 2002). After optimization of the isolation procedures, many different kinds of cells from a variety of plant tissues and species were successfully isolated (Kerk et al., 2003; Nakazono et al., 2003). The most important application of LMD in researching sexual plant reproduction is to isolate specific cells of female gametes and gametophytes, as well as embryos, and helps to overcome the difficulty of accessing specific cells from the complex maternal tissues. With the LMD-harvested pure cells, the transcriptomes of specific cells were investigated and provided sufficient information for gene expression in female

gametes, zygotes and embryos (Casson et al., 2005; Le et al., 2007; Wuest et al., 2010). The LMD technique has the advantages of ease of use and automating the tedious procedures of cell collection. However, it is still very labor intensive in the sample preparation steps and it is easy to contaminate samples with neighboring cells, especially for small or few layers of cells and tissues such as in meristem.

Fluorescence activated cell sorting

The fluorescence activated cell sorting (FACS) was developed for quick and accurate isolation of specific cells. FACS is the technique of choice to purify specific cell populations, and allows the purification of individual cells based on size, granularity and fluorescence. The purification process requires a flow cytometer with sorting capacity and appropriate software. In the process of the sorting, suspension cells are passed as a stream in droplets with each containing a single cell in front of a laser. The fluorescence detection system detects the cells of interest, based on predetermined fluorescent parameters. The instrument first applies a charge to the droplet containing a cell of interest, and then an electrostatic deflection system facilitates collection of charged droplets into appropriate collection tubes (Givan, 2001). There are two common approaches to label specific cells with fluorescence, one is cell pre-staining with fluorescently-tagged monoclonal antibodies that recognize specific surface-markers on the desired cell population (Herzenberg and De Rosa, 2000), and the other is genetic modification of the cells of interest with fluorescent protein driven by cell-specific promoters (Brandt et al., 1999).

The first application of FACS in plant cell-specific transcriptome analyses was reported in the sperm cells of maize (Engel et al., 2003). Thereafter, the method was used to isolate sperm cells of *Arabidopsis* for transcriptome analysis at a genome-wide level (Borges et al., 2008). Apart from sperm cells, FACS was adopted to isolate different types of protoplasts from *Arabidopsis* root, which provided materials for obtaining detailed spatiotemporal gene expression maps of the root (Birnbaum et al., 2003; Brady et al., 2007). Although FACS requires specialized equipment and training, it is the preferred method for isolating highly purified cell populations. Up to now, the application of FACS in plant cells has been relatively limited since it relies on the prior identification of cell-specific markers; however, the recent availability of increasingly cell type-specific fluorescent reporter lines will greatly advance the progress of FACS application.

Techniques of cell type-specific transcriptome analysis

cDNA library construction from limited cells

In higher eukaryotes, biological processes such as cell growth and differentiation are regulated by cell-specific gene expression. To understand the mechanism of molecular regulation, the transcriptomes for special types of cells should be studied to provide an overview of their expression profiles. The construction of cDNA libraries for specific cells serves as a useful method for decades. However, normal cDNA library construction is usually insufficient for isolating rare RNA at the ng-level from limited cell samples. To overcome the limitation of cell population, PCR-based cDNA library construction was developed to generate cDNA libraries at the level of a few cells, and the pre-amplification of the first-strand cDNA by long-distance PCR provided the necessary amount of cDNA for library construction (Belyavsky et al., 1989; Dresselhaus et al., 1994). The wide existence of abundant species of cDNA hampers the identification of genes with low abundance. Subtractive cDNA hybridization, a kinetic method to deplete high abundance cDNA between two cell samples, is a powerful approach to identify and isolate differentially expressed genes. The subtracted cDNA libraries greatly enrich the differentially expressed sequences, and enable identification of cDNA corresponding to species of mRNA present at very low levels or only in a small fraction of cells from complex tissues (Diatchenko et al., 1996; Lê et al., 2005; Ning et al., 2006; Hu et al., 2010). These specialized cDNA libraries from a variety of cell types are a necessary prerequisite to more detailed analysis of the distribution of mRNA populations (Sprunck et al., 2005; Okada et al., 2006; Yang et al., 2006). Further comparison of library transcripts derived from different cell types leads to ease of isolation of the expressed sequences unique to specific cell-types (Sprunck et al., 2005; Gou et al., 2009). cDNA library construction provides an approach for investigating the transcriptome profiles of different cells, but the information obtained greatly depends on the sequencing of library transcripts, which is often time consuming and expensive.

DNA microarray

Apart from the cDNA library, the DNA microarray technique is another important tool in the research of functional genomics, and can help researchers bypass cDNA library construction and avoid bias generation in the sub-cloning process. The advantage of this method lies in miniaturization, automation and parallelism, and it permits large-scale and genome-wide acquisition of quantitative transcriptome information from different cell samples. At present, there are two main approaches in DNA microarray fabrication, involving *in situ* synthesis of oligonucleotides (oligonucleotide microarrays) and deposition of pre-synthesized DNA fragments (cDNA microarrays) on solid surfaces (Aharoni and Vorst, 2002). DNA microarray technology has been widely used in functional genomics, disease genomics, systems biology and other fields, with the main applications in comprehensive, simultaneous gene expression monitoring and in DNA

variation analyses for the identification of mutations and genotyping of polymorphisms. For plant cells, the DNA microarray technique potentially allows a global analysis of gene expression at whole-genome level, and has been used to compare transcriptional changes in different types of cells. These cell-specific transcriptome analyses provide important insights into the regulation networks that are linked to cellular identities and functions (Borges et al., 2008; Haerizadeh et al., 2009; Le et al., 2010; Wei et al., 2010). Exciting new information will be gained in the years to come, not only from genome-wide expression analyses on a few model plant species, but also from extensive application to less thoroughly studied species on a relatively limited scale (Okada et al., 2007; Gou et al., 2009). Since the DNA microarray has already become a mature systemic platform, there are standardized protocols for sample preparation, experiment execution and data analysis, along with large public databases available to all scientists. However, the shortcoming of DNA microarray is that it is a 'closed system', which can only detect the characteristics of known sequences (or with limited variation) in a large number of biological samples, and rare RNA expressed at low levels is usually beyond its detection capability.

RNA-Seq by NG sequencing

Since its invention, DNA sequencing has played an essential role in the advancement of molecular biology. The first-generation sequencing technique was based on the fluorescently labeled Sanger's chain-termination method developed in the mid-1970s (Sanger, 1988). The NG sequencing technology, which depends on massive parallel DNA sequencing platforms, greatly increased analysis speed and reduced sequencing costs (Shendure and Ji, 2008). Three widely used NG commercial platforms (Illumina Genome Analyzer, Roche 454 Genome Sequencer and Life Technologies SOLiD System) were all invented and developed toward the end of 1990s, and commercialized in the middle of the first decade of the century. All utilize the sequencing-by-synthesis strategy, despite their variation in DNA array formation, fragment amplification, enzyme-based sequencing biochemistry and subsequent sequential analysis (Ansorge, 2009). Next-generation DNA sequencing, which dramatically accelerated biologic and biomedical research by making the comprehensive analysis of genomes inexpensive, and enabling it become routine and widespread, is also starting to be used for transcriptome analysis of plant cells. Comparative research between traditional capillary sequencing and NG ultra high-throughput techniques was carried out in *Arabidopsis*, *Eschscholzia californica* and *Persea americana*, and showed that NG sequencing was a dramatic advance over capillary-based sequencing in terms of sequence coverage (Wall et al., 2009). In recent reports, NG sequencing has been used for comprehensive studies of transcriptomes at single-base resolution in rice (Zhang et al., 2010), and investigation of DNA methylation, activating and repressive histone modifications, small RNAs and mRNA in the maize genome (Elling and Deng, 2009). Moreover, cell-specific digital gene expression profiling with only a single mouse blastomere was successfully carried out using mRNA-Seq assay ~75% more expressed genes were detected with this technique compared to microarray, and this assay identified 1753 previously unknown splice junctions (Tang et al., 2009). Although microarray is still the most widely used technique in current reports of plant cell transcriptome studies, RNA-Seq by NG sequencing has already displayed extraordinary charisma in all areas of genome research, and it is increasingly replacing DNA microarrays in cell-specific transcriptome analysis.

Application of the techniques in plant sexual reproduction

Cell type-specific transcriptome analysis has been carried out on many different kinds of cells in higher plants. Here we focus on the application in gene expression analysis of male and female gametes, zygotes and early proembryos, and summarize some recent advances in sexual plant reproduction.

Male gametes

A pollen grain is the highly reduced haploid male gametophyte in flowering plants, and consists of just two or three cells when it matures and is released from the anther. Mature pollen grains are carried by wind or insects onto the stigma of the pistil, where the pollen germinates and produces pollen tubes. Then the pollen tubes grow through the style into the female gametophyte in the ovule. In the tips of pollen tubes, two sperm cells are delivered into the embryo sac and fuse with an egg cell and a central cell, respectively. This double fertilization event along with the functional specialization of the male gametophyte is considered to be a key innovation in the evolutionary success of flowering plants. During recent decades, there have been great advances in understanding the molecular regulation network of male gamete and gametophyte development. The first cDNA libraries of male gametes were constructed in the mature pollen of maize and *Tradescantia. paludosa*, and further northern hybridizations revealed the existence of pollen-specific cDNA clones (Stinson et al., 1987). In recent years, several cDNA libraries from mature anthers of tomato, pollen of *Brassica campestris*, and generative cells of lily were constructed, and library screenings with northern blot analysis led to the isolation of male-gamete-specific genes, such as *LAT52* and *LGC1* (Twell et al., 1989; Theerakulpisut et al., 1991; Xu et al., 1998, 1999). Since large scale EST (expressed sequence tag) sequencing was not applied in these studies, the information obtained about gene expression in the male gametes was still relatively limited.

The wide application of microarray and large-scale cDNA library sequencing in the first decade of this century greatly promoted the characterization of transcriptomes in male gametes, such as pollen development at different stages in *Arabidopsis* (Honys and Twell, 2004), rapeseed (Whittle et al. 2010) and rice (Hobo et al., 2008; Wei et al., 2010). These provide an overview of the dynamic gene expression in male gametes, revealing a large-scale repression of early program genes and the activation of a unique late gene-expression in the progress of pollen maturation. More researches, based on microarray analysis, compared the transcriptional profiles of mature/hydrated pollens with sporophytic tissues in the tri-cellular pollen model plant *Arabidopsis* (Becker et al., 2003; Honys and Twell, 2003; Pina et al., 2005) and the bi-cellular pollen plant soybean (Haerizadeh et al., 2009). All these studies indicated that the pollen possesses a restricted and unique repertoire of genes, with a significantly greater proportion of specifically expressed genes compared to sporophyte tissues. Moreover, comparison of the upregulated transcription factors between soybean and *Arabidopsis* reveals some divergence in the numbers and kinds of regulatory proteins in both species.

Considering the relative complexity of the content in pollen, the transcriptomes of generative cells in lily were studied by EST analysis of cDNA library and microarrays (Okada et al., 2006, 2007). The results showed that a subset of genes was specifically expressed in generative cells, and that some genes showed significant similarities to the male gametophyte-specific genes from maize and *Arabidopsis*, indicating conservation of such genes across plant species. Using the technique of directly isolating sperm cells, the first cDNA library from the sperm cells of rice was constructed (Gou et al., 2001), and the sperm cell cDNA libraries of tobacco (Xu et al., 2002; Xin et al., 2010) and maize (Engel et al., 2003) were also obtained. Screening and sequencing of the sperm-cell-specific libraries enabled the identification of the special sperm-cell-expressed transcripts, and revealed the complex and special composition of mRNAs in the sperm. In the isolated sperm cells of *Arabidopsis* using FACS, the genome-scale transcriptome analysis with the GeneChip technique revealed that sperm cells possessed distinct transcriptional information compared with those of seedlings and also pollen (Borges et al., 2008). To study the dimorphism of sperm cells, two populations of the dimorphic sperm cells in *P. zeylanica* were collected to construct cDNA libraries and carry out microarray analysis – this revealed the distinct complements of expressed transcripts in the two kinds of sperm cells (Gou et al., 2009). All these analyses indicate that male gametes possess a unique composition of gene expression compared with sporophytic tissues, and that a significant proportion of transcripts displayed a male gamete cell-specific expression. The identification and functional characterization of the genes promises the elucidation of their regulatory networks operating in male gamete development and double fertilization of higher plants.

Female gametes

In flowering plants, the haploid female gametophyte is a typical seven-celled structure with four types of cells: an egg cell, a central cell, two synergid cells, and three antipodal cells. These cells derive from a functional megaspore and perform their respective essential functions required for double fertilization and seed development. Differentiation of these distinct cells likely involves the coordinated changes of gene expression regulated by transcription factors. Therefore, understanding of female gametophyte cell differentiation and function is required to study the gene regulatory networks operating in each cell type. For this purpose, many researches have been conducted to elucidate the molecular mechanisms of cell specification and differentiation in the female gametes and gametophytes. The first representative cDNA library of egg cells in higher plants was generated from 128 isolated maize egg-cells using a reverse transcriptase/polymerase chain reaction (RT-PCR) method (Dresselhaus et al., 1994). *Egg Apparatus 1* of maize from this library was identified and proved to function in guiding pollen tubes into the micropyle (Márton et al., 2005). A more detailed transcriptome assay was carried out with cDNA libraries from female gametophytes and egg cells of maize (Yang et al., 2006), wheat (Kumlehn et al., 2001; Sprunck et al., 2005) and tobacco (Ning et al., 2006), and more ESTs were sequenced and analyzed for investigating transcription profiles in female gametes and for isolating the egg cell-specific genes.

For identifying specific and differentially expressed transcripts in different gamete cells, the subtractive cDNA libraries between maize egg cells and central cells were constructed, and 340 differentially expressed clones were identified through microarray screening (Lê et al., 2005). Taking advantage of the protruding embryo sacs and easily obtained member cells of *T. fournieri*, the cDNA library from manually-isolated synergid cells was constructed, and then several cysteine-rich polypeptides, which function as pollen tube attractants secreted from synergid cells, were successfully identified (Okuda et al., 2009). To determine cell-type-specific expression profiles in the female gametophyte of *Arabidopsis*, the LMD technique was used to isolate individual cells including the synergids and the two female gametes (egg and central cells) as samples for microarray analysis, and revealed that these distinct types of cells were characterized by their differentially expressed post-transcriptional regulatory modules and metabolic pathways (Wuest et al., 2010).

Besides direct analysis of the specific gamete cells, another widely applied approach to characterize genes involved in female gamete development is based on analysis in the ovules. Comparative transcriptome analyses between the normal ovules and those lacking the complete embryo sacs such as the *determinate infertile1* (*dif1*), *sporocyteless* (*spl*) and *coatlique* (*coa*) mutant ovules were performed by several different groups with the GeneChip hybridization (Yu et al.,

2005; Johnston et al., 2007; Steffen et al., 2007). Similar research was carried out by large-scale real-time RT-PCR screening or high-throughput RNA-Seq and whole-genome tiling arrays in the *dif1* ovules and *myb98* ovules that were impaired in pollen tube attraction (Jones-Rhoades et al., 2007; Wang et al., 2010). Apart from comparative analysis with mutants, the cDNA libraries of rapeseed ovules were also constructed and 10468 ESTs obtained, providing new insights into the transcriptional profile of rarely studied ovules (Whittle et al., 2010). The results of the analyses defined a large number of embryo sac-expressed transcripts, and some candidate transcripts were found to be expressed in specific cells of the female gametophytes. Since these specific transcripts of gamete cells may act during embryo sac development, fertilization and even embryogenesis, further functional characterization of these genes will promote elucidating the molecular mechanisms and regulatory networks in the processes of development.

Zygotes and embryos

Embryogenesis is a crucial developmental event in the life cycle of flowering plants. During embryogenesis, a single-celled zygote follows a defined pattern of cell division and differentiation to form a mature embryo. To investigate the developmental expression of mRNAs in embryos, the isolated embryos from many species (e.g., soybean, rapeseed, barley and rice) were first used for the cDNA library construction and the isolation of embryo specific transcripts (Goldberg et al., 1981; Loader et al., 1993; Ranford et al., 2002; Ito et al., 2004). Hand-dissected embryo proper and suspensor cells from the globular embryos of the scarlet runner bean were used for differential screening to identify suspensor-specific mRNAs (Weterings et al., 2001). In maize and rice, the expressed genes during the embryo development were investigated by DNA microarray and EST sequencing (Lee et al., 2002; Ge et al., 2008), providing more detailed molecular insights into embryonic development. For more specific analysis of specific cells of embryos, LMD was applied to isolate the apical and basal domains of the globular-, heart- and torpedo-stage embryos from tissue sections of *Arabidopsis* (Casson et al., 2005; Spencer et al., 2007). In the development of embryos, the transition from globular embryo to heart embryo is a key phase of polarity establishment and organ differentiation. Global transcriptional profiling with specific domains of embryos not only enabled the isolation of embryo cell type-specific genes, but also revealed the dynamics and complexity of gene expression of the embryos in both spatial and temporal dimensions, which provided much useful information for understanding of embryo polarity establishment and organ differentiation.

For investigating gene expression in zygotes and early embryos, a subtractive cDNA library was constructed to compare the gene expression profiles between isolated egg cells and zygotes of tobacco. The identification of many zygote-expressed genes revealed that zygotic gene activation occurs shortly after fertilization in higher plants (Ning et al., 2006). The first division of zygotes is a critical event during early embryogenesis, and thereafter, the apical and basal daughter cells become specified to follow their different developmental pathways. To gain understanding of the processes, the transcriptional profiles of the two-celled proembryos in wheat (Sprunck et al., 2005), the apical and basal cells of *in vitro* two-celled proembryos in maize (Okamoto et al., 2005) and the *in vivo* apical and basal cells in tobacco (Hu et al., 2010) were obtained and investigated. Using an *in vitro* zygote culture system, the two-celled proembryos with the same size of cells from symmetric division of zygotes were obtained as a comparison sample. The differential gene expression in the two kinds of different zygotic division processes, the asymmetric division *in vivo* and symmetric division *in vitro*, was well studied (Hu et al. unpublished data). The analyses showed a global view of the transcript profiles in the different two-celled proembryos, and the apical and basal cells, respectively. The identification of transcripts with specific expressions, which were up- or downregulated only in the apical or basal cells and in the asymmetric/symmetric two-celled proembryos, will provide important clues for understanding of zygote activation and division, and proembryo polarity formation and differentiation.

To identify gene expressions in the specific stages of seed development, LMD was also applied to dissect different compartments of the entire seeds (e.g., endosperm, suspensor, embryo proper, endothelium, inner integument, outer integument, epidermis and hilum from a globular-stage soybean seed). The global gene-expression profiles in these compartments were all investigated using microarray-based analysis, and approximately 20000 diverse transcripts were found in the whole-mount globular-stage soybean seeds, with small sets of seed region-specific mRNAs (Le et al., 2007). In *Arabidopsis*, the transcription profiles of ovules or seeds from fertilization to maturation and germination were obtained and studied, including ovules with an unfertilized egg cell, zygotes, globular embryos, cotyledon embryos, mature green embryos, postmature green embryos and postgermination seedlings. The amount of expressed unique mRNAs varied from about 9000 to 14000 at different developmental stages with about 7000 mRNAs detected at all stages of tested seeds (Le et al., 2010). This research provided a detailed map of gene expression in various seed compartments at different developmental stages and enabled the identification of embryo, endosperm and seed coat-specific transcripts. These ovules or seed-specific genes should facilitate the elucidation of regulatory networks for programming seed development.

Future perspectives

Although the cell-specific transcriptome analysis reviewed

here provides much information about gene expression in the gamete cells, zygotes and embryos, and has greatly enhanced understanding of the molecular mechanisms involved in regulating the sexual reproduction process, many challenges still remain. The first challenge is that it is currently very difficult to elucidate the regulation network controlling gametogenesis and embryogenesis. Although transcriptome information of angiosperm gametes and embryos has greatly increased in recent years, there is still a need for larger-scale, more specific and extensive genome-wide studies, especially for processes of zygote activation and directional division, and embryogenesis and early polar axis establishment. The application of FACS and RNA-Seq by NG sequencing will accelerate genome-wide detection of mRNA expression levels in the sexual reproduction process. There is still much work remaining to elucidate the detailed functions of cell-specific transcripts. Another challenge is whether the expressed mRNAs indeed translate to functional proteins. Proteomic approaches are useful tools for investigating transcriptional networks at the cellular level. Improved high-throughput proteomics techniques have shifted attention to protein profiling, which attempts to identify all proteins present in particular tissues or cells (Baginsky and Gruissem, 2006). A maize proteome map has been established in isolated egg cells (Okamoto et al., 2004), and more sensitive methods for cell-specific proteomic analysis will hopefully be established in the near future. Despite the challenges, the recent development in cell type-specific transcriptome and proteome analysis techniques holds great promise for the future. Analysis of transcriptomes and proteomes, and further identification of gene functions and signaling pathways in the gametes and embryos, will ultimately lead to the identification of the molecular networks that regulate the processes of sexual plant reproduction.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 30970277) and the Major State Basic Research Program of China (No. 2007CB108704)

References

Aharoni A, Vorst O (2002). DNA microarrays for functional plant genomics. Plant Mol Biol, 48(1–2): 99–118

Ansorge W J (2009). Next-generation DNA sequencing techniques. New Biotechnol, 25(4): 195–203

Asano T, Masumura T, Kusano H, Kikuchi S, Kurita A, Shimada H, Kadowaki K (2002). Construction of a specialized cDNA library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem. Plant J, 32(3): 401–408

Baginsky S, Gruissem W (2006). *Arabidopsis thaliana* proteomics: from proteome to genome. J Exp Bot, 57(7): 1485–1491

Becker J D, Boavida L C, Carneiro J, Haury M, Feijó J A (2003). Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. Plant Physiol, 133(2): 713–725

Belyavsky A, Vinogradova T, Rajewsky K (1989). PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. Nucleic Acids Res, 17(8): 2919–2932

Birnbaum K, Shasha D E, Wang J Y, Jung J W, Lambert G M, Galbraith D W, Benfey P N (2003). A gene expression map of the *Arabidopsis* root. Science, 302(5652): 1956–1960

Borges F, Gomes G, Gardner R, Moreno N, McCormick S, Feijó J A, Becker J D (2008). Comparative transcriptomics of *Arabidopsis* sperm cells. Plant Physiol, 148(2): 1168–1181

Brady S M, Orlando D A, Lee J Y, Wang J Y, Koch J, Dinneny J R, Mace D, Ohler U, Benfey P N (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. Science, 318(5851): 801–806

Brandt S, Kehr J, Walz C, Imlau A, Willmitzer L, Fisahn J (1999). Technical advance: A rapid method for detection of plant gene transcripts from single epidermal, mesophyll and companion cells of intact leaves. Plant J, 20(2): 245–250

Brandt S, Kloska S, Altmann T, Kehr J (2002). Using array hybridization to monitor gene expression at the single cell level. J Exp Bot, 53 (379): 2315–2323

Cao Y, Reece A, Russell S D (1996). Isolation of viable sperm cells from tobacco (*Nicotiana tabacum*). Zygote, 4(2): 81–84

Casson S, Spencer M, Walker K, Lindsey K (2005). Laser capture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis.* Plant J, 42(1): 111–123

Chen S H, Yang Y H, Liao J P, Kuang A X, Tian H Q (2008). Isolation of egg cells and zygotes of *Torenia fournieri* L. and determination of their surface charge. Zygote, 16(2): 179–186

Diatchenko L, Lau Y F, Campbell A P, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov E D, Siebert P D (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci USA, 93(12): 6025–6030

Dresselhaus T, Lörz H, Kranz E (1994). Representative cDNA libraries from few plant cells. Plant J, 5(4): 605–610

Elling A A, Deng X W (2009). Next-generation sequencing reveals complex relationships between the epigenome and transcriptome in maize. Plant Signal Behav, 4(8): 760–762

Emmert-Buck M R, Bonner R F, Smith P D, Chuaqui R F, Zhuang Z, Goldstein S R, Weiss R A, Liotta L A (1996). Laser capture microdissection. Science, 274(5289): 998–1001

Engel M L, Chaboud A, Dumas C, McCormick S (2003). Sperm cells of *Zea mays* have a complex complement of mRNAs. Plant J, 34(5): 697–707

Fu C M, Sun M X, Zhou C, Yang H Y (1996). Isolation of fertilized embryo sacs and zygotes and triggering of zygote division *in vitro* in *Nicotiana tabacum.* Acta Bot Sin, 38: 262–267

Ge X, Chen W, Song S, Wang W, Hu S, Yu J (2008). Transcriptomic profiling of mature embryo from an elite super-hybrid rice LYP9 and its parental lines. BMC Plant Biol, 8(1): 114

Givan A L (2001). Flow Cytometry First Principles. New York: Wiley-Liss,. 273

Goldberg R B, Hoschek G, Ditta G S, Breidenbach R W (1981).

Developmental regulation of cloned superabundant embryo mRNAs in soybean. Dev Biol, 83(2): 218–231

Gou X, Yuan T, Wei X, Russell S D (2009). Gene expression in the dimorphic sperm cells of *Plumbago zeylanica*: transcript profiling, diversity, and relationship to cell type. Plant J, 60(1): 33–47

Gou X P, Xu Y, Tang L, Yan F, Chen F (2001). Representative cDNA library from isolated rice sperm cells. Acta Bot Sin, 43: 1093–1096

Haerizadeh F, Wong C E, Bhalla P L, Gresshoff P M, Singh M B (2009). Genomic expression profiling of mature soybean (*Glycine max*) pollen. BMC Plant Biol, 9(1): 25

Herzenberg L A, De Rosa S C (2000). Monoclonal antibodies and the FACS: complementary tools for immunobiology and medicine. Immunol Today, 21(8): 383–390

Hobo T, Suwabe K, Aya K, Suzuki G, Yano K, Ishimizu T, Fujita M, Kikuchi S, Hamada K, Miyano M, Fujioka T, Kaneko F, Kazama T, Mizuta Y, Takahashi H, Shiono K, Nakazono M, Tsutsumi N, Nagamura Y, Kurata N, Watanabe M, Matsuoka M (2008). Various spatiotemporal expression profiles of anther-expressed genes in rice. Plant Cell Physiol, 49(10): 1417–1428

Holm P B, Knudsen S, Mouritzen P, Negri D, Olsen F L, Roue C (1994). Regeneration of fertile barley plants from mechanically isolated protoplasts of the fertilized egg cell. Plant Cell, 6(4): 531–543

Honys D, Twell D (2003). Comparative analysis of the *Arabidopsis* pollen transcriptome. Plant Physiol, 132(2): 640–652

Honys D, Twell D (2004). Transcriptome analysis of haploid male gametophyte development in *Arabidopsis.* Genome Biol, 5(11): R85

Hu S Y, Li L G, Zhu C (1985). Isolation of viable embryo sacs and their protoplasts of *Nicotiana tabacum.* Acta Bot Sin, 27: 337–344

Hu S Y, Yang H Y (2002). Biology of Angiosperm Fertilization.Beijing: Science Press, 250

Hu T X, Yu M, Zhao J (2010). Comparative transcriptional profiling analysis of the two daughter cells from tobacco zygote reveals the transcriptome differences in the apical and basal cells. BMC Plant Biol, 10(1): 167

Huang B Q, Russell S D (1989). Isolation of fixed and viable eggs, central cells, and embryo sacs from ovules of *Plumbago zeylanica.* Plant Physiol, 90(1): 9–12

Ito Y, Chujo A, Eiguchi M, Kurata N (2004). Radial axis differentiation in a globular embryo is marked by HAZ1, a PHD-finger homeobox gene of rice. Gene, 331: 9–15

Johnston A J, Meier P, Gheyselinck J, Wuest S E, Federer M, Schlagenhauf E, Becker J D, Grossniklaus U (2007). Genetic subtraction profiling identifies genes essential for *Arabidopsis* reproduction and reveals interaction between the female gametophyte and the maternal sporophyte. Genome Biol, 8(10): R204

Jones-Rhoades M W, Borevitz J O, Preuss D (2007). Genome-wide expression profiling of the *Arabidopsis* female gametophyte identifies families of small, secreted proteins. PLoS Genet, 3(10): 1848–1861

Karrer E E, Lincoln J E, Hogenhout S, Bennett A B, Bostock R M, Martineau B, Lucas W J, Gilchrist D G, Alexander D (1995). *In situ* isolation of mRNA from individual plant cells: creation of cell-specific cDNA libraries. Proc Natl Acad Sci U S A, 92(9): 3814–3818

Katoh N, Lörz H, Kranz E (1997). Isolation of viable egg cells of rape (*Brassica napus* L.). Zygote, 5(1): 31–33

Kerk N M, Ceserani T, Tausta S L, Sussex I M, Nelson T M (2003). Laser capture microdissection of cells from plant tissues. Plant Physiol, 132(1): 27–35

Kranz E, Bautor J, Lörz H (1991). *In vitro* fertilization of single, isolated gametes of maize mediated by electrofusion. Sex Plant Reprod, 4: 12–16

Kumlehn J, Lörz H, Kranz E (1998). Differentiation of isolated wheat zygotes into embryos and normal plants. Planta, 205(3): 327–333

Kumlehn K, Czihal A, Matzk L (2001). Parthenogenetic egg cells of wheat: cellular and molecular studies. Sex Plant Reprod, 14(4): 239–243

Le B H, Cheng C, Bui A Q, Wagmaister J A, Henry K F, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, Drews G N, Fischer R L, Okamuro J K, Harada J J, Goldberg R B (2010). Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. Proc Natl Acad Sci USA, 107(18): 8063–8070

Le B H, Wagmaister J A, Kawashima T, Bui A Q, Harada J J, Goldberg R B (2007). Using genomics to study legume seed development. Plant Physiol, 144(2): 562–574

Lê Q, Gutièrrez-Marcos J F, Costa L M, Meyer S, Dickinson H G, Lörz H, Kranz E, Scholten S (2005). Construction and screening of subtracted cDNA libraries from limited populations of plant cells: a comparative analysis of gene expression between maize egg cells and central cells. Plant J, 44(1): 167–178

Lee J M, Williams M E, Tingey S V, Rafalski J A (2002). DNA array profiling of gene expression changes during maize embryo development. Funct Integr Genomics, 2(1–2): 13–27

Loader N M, Woolner E M, Hellyer A, Slabas A R, Safford R (1993). Isolation and characterization of two *Brassica napus* embryo acyl-ACP thioesterase cDNA clones. Plant Mol Biol, 23(4): 769–778

Márton M L, Cordts S, Broadhvest J, Dresselhaus T (2005). Micropylar pollen tube guidance by egg apparatus 1 of maize. Science, 307 (5709): 573–576

Nakazono M, Qiu F, Borsuk L A, Schnable P S (2003). Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. Plant Cell, 15(3): 583–596

Nelson T, Tausta S L, Gandotra N, Liu T (2006). Laser microdissection of plant tissue: what you see is what you get. Annu Rev Plant Biol, 57 (1): 181–201

Ning J, Peng X B, Qu L H, Xin H P, Yan T T, Sun M (2006). Differential gene expression in egg cells and zygotes suggests that the transcriptome is restructed before the first zygotic division in tobacco. FEBS Lett, 580(7): 1747–1752

Okada T, Bhalla P L, Singh M B (2006). Expressed sequence tag analysis of *Lilium longiflorum* generative cells. Plant Cell Physiol, 47 (6): 698–705

Okada T, Singh M B, Bhalla P L (2007). Transcriptome profiling of *Lilium longiflorum* generative cells by cDNA microarray. Plant Cell Rep, 26(7): 1045–1052

Okamoto T, Higuchi K, Shinkawa T, Isobe T, Lörz H, Koshiba T, Kranz E (2004). Identification of major proteins in maize egg cells. Plant Cell Physiol, 45(10): 1406–1412

Okamoto T, Scholten S, Lörz H, Kranz E (2005). Identification of genes that are up- or down-regulated in the apical or basal cell of

maize two-celled embryos and monitoring their expression during zygote development by a cell manipulation- and PCR-based approach. Plant Cell Physiol, 46(2): 332–338

Okuda S, Tsutsui H, Shiina K, Sprunck S, Takeuchi H, Yui R, Kasahara R D, Hamamura Y, Mizukami A, Susaki D, Kawano N, Sakakibara T, Namiki S, Itoh K, Otsuka K, Matsuzaki M, Nozaki H, Kuroiwa T, Nakano A, Kanaoka M M, Dresselhaus T, Sasaki N, Higashiyama T (2009). Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature, 458(7236): 357–361

Pina C, Pinto F, Feijó J A, Becker J D (2005). Gene family analysis of the *Arabidopsis* pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. Plant Physiol, 138(2): 744–756

Ranford J C, Bryce J H, Morris P C (2002). PM19, a barley (*Hordeum vulgare* L.) gene encoding a putative plasma membrane protein, is expressed during embryo development and dormancy. J Exp Bot, 53 (366): 147–148

Russell S D (1986). Isolation of sperm cells from the pollen of *Plumbago zeylanica.* Plant Physiol, 81(1): 317–319

Russell S D (1992). Double fertilization. Int Rev Cytol, 40: 357–390

Sanger F (1988). Sequences, sequences, and sequences. Annu Rev Biochem, 57(1): 1–28

Shendure J, Ji H (2008). Next-generation DNA sequencing. Nat Biotechnol, 26(10): 1135–1145

Simone N L, Bonner R F, Gillespie J W, Emmert-Buck M R, Liotta L A (1998). Laser-capture microdissection: opening the microscopic frontier to molecular analysis. Trends Genet, 14(7): 272–276

Spencer M W, Casson S A, Lindsey K (2007). Transcriptional profiling of the *Arabidopsis* embryo. Plant Physiol, 143(2): 924–940

Sprunck S, Baumann U, Edwards K, Langridge P, Dresselhaus T (2005). The transcript composition of egg cells changes significantly following fertilization in wheat (*Triticum aestivum* L.). Plant J, 41 (5): 660–672

Steffen J G, Kang I H, Macfarlane J, Drews G N (2007). Identification of genes expressed in the *Arabidopsis* female gametophyte. Plant J, 51 (2): 281–292

Stinson J R, Eisenberg A J, Willing R P, Pe M E, Hanson D D, Mascarenhas J P (1987). Genes expressed in the male gametophyte of flowering plants and their isolation. Plant Physiol, 83(2): 442–447

Tanaka I (1988). Isolation of generative cells and their protoplasts from pollen of *Lilium longiflorum*. Protoplasma, 142(1): 68–73

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch B B, Siddiqui A, Lao K, Surani M A (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods, 6(5): 377–382

Theerakulpisut P, Xu H, Singh M B, Pettitt J M, Knox R B (1991). Isolation and developmental expression of Bcp1, an anther-specific cDNA clone in *Brassica campestris.* Plant Cell, 3(10): 1073–1084

Twell D, Wing R, Yamaguchi J, McCormick S (1989). Isolation and expression of an anther-specific gene from tomato. Mol Gen Genet, 217(2–3): 240–245

Wall P K, Leebens-Mack J, Chanderbali A S, Barakat A, Wolcott E, Liang H, Landherr L, Tomsho L P, Hu Y, Carlson J E, Ma H, Schuster S C, Soltis D E, Soltis P S, Altman N, dePamphilis C W (2009). Comparison of next generation sequencing technologies for transcriptome characterization. BMC Genomics, 10(1): 347

Wang D, Zhang C, Hearn D J, Kang I H, Punwani J A, Skaggs M I, Drews G N, Schumaker K S, Yadegari R (2010). Identification of transcription-factor genes expressed in the *Arabidopsis* female gametophyte. BMC Plant Biol, 10(1): 110

Wang D Y, Zhang Q, Liu Y, Lin Z F, Zhang S X, Sun M X, Sodmergen (2010). The levels of male gametic mitochondrial DNA are highly regulated in angiosperms with regard to mitochondrial inheritance. Plant Cell, 22(7): 2402–2416

Wei L Q, Xu W Y, Deng Z Y, Su Z, Xue Y, Wang T (2010). Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in *Oryza sativa.* BMC Genomics, 11(1): 338

Weterings K, Apuya N R, Bi Y, Fischer R L, Harada J J, Goldberg R B (2001). Regional localization of suspensor mRNAs during early embryo development. Plant Cell, 13(11): 2409–2425

Whittle C A, Malik M R, Li R, Krochko J E (2010). Comparative transcript analyses of the ovule, microspore, and mature pollen in *Brassica napus.* Plant Mol Biol, 72(3): 279–299

Wuest S E, Vijverberg K, Schmidt A, Weiss M, Gheyselinck J, Lohr M, Wellmer F, Rahnenführer J, von Mering C, Grossniklaus U (2010). *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. Curr Biol, 20(6): 506–512

Xin H P, Peng X B, Ning J, Yan T T, Ma L G, Sun M X (2010). Expressed sequence-tag analysis of tobacco sperm cells reveals a unique transcriptional profile and selective persistence of paternal transcripts after fertilization. Sex Plant Reprod,

Xu H, Swoboda I, Bhalla P L, Sijbers A M, Zhao C, Ong E K, Hoeijmakers J H, Singh M B (1998). Plant homologue of human excision repair gene ERCC1 points to conservation of DNA repair mechanisms. Plant J, 13(6): 823–829

Xu H, Swoboda I, Bhalla P L, Singh M B (1999). Male gametic cell-specific gene expression in flowering plants. Proc Natl Acad Sci USA, 96(5): 2554–2558

Xu H, Weterings K, Vriezen W, Feron R, Xue Y, Derksen J, Mariani C (2002). Isolation and characterization of male-germ-cell transcripts in *Nicotiana tabacum.* Sex Plant Reprod, 14(6): 339–346

Yang H, Kaur N, Kiriakopolos S, McCormick S (2006). EST generation and analyses towards identifying female gametophyte-specific genes in *Zea mays* L. Planta, 224(5): 1004–1014

Yu H J, Hogan P, Sundaresan V (2005). Analysis of the female gametophyte transcriptome of *Arabidopsi*s by comparative expression profiling. Plant Physiol, 139(4): 1853–1869

Zhang G, Guo G, Hu X, Zhang Y, Li Q, Li R, Zhuang R, Lu Z, He Z, Fang X, Chen L, Tian W, Tao Y, Kristiansen K, Zhang X, Li S, Yang H, Wang J, Wang J (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Res, 20(5): 646–654

Zhang Z, Xu H, Singh M B, Russell S D (1998). Isolation and collection of two populations of viable sperm cells from the pollen of *Plumbago zeylanica.* Zygote, 6(4): 295–298

Zhao J, Zhou C, Yang H Y (2000). Isolation and *in vitro* culture of zygotes and central cells of *Oryza sativa* L. Plant Cell Rep, 19(3): 321–326