Mining meiosis and gametogenesis with DNA microarrays

Ulrich Schlecht and Michael Primig*

Gametogenesis is a key developmental process that involves complex transcriptional regulation of numerous genes including many that are conserved between unicellular eukaryotes and mammals. Recent expression-profiling experiments using microarrays have provided insight into the co-ordinated transcription of several hundred genes during mitotic growth and meiotic development in budding and fission yeast. Furthermore, microarray-based studies have identified numerous loci that are regulated during the cell cycle or expressed in a germ-cell specific manner in eukaryotic model systems like *Caenorhabditis elegans*, *Mus musculus* as well as *Homo sapiens*. The unprecedented amount of information produced by post-genome biology has spawned novel approaches to organizing biological knowledge using currently available information technology. This review outlines experiments that contribute to an emerging comprehensive picture of the molecular machinery governing sexual reproduction in eukaryotes.

During meiosis a competent diploid cell replicates its DNA once and then undergoes two consecutive divisions followed by haploid gamete differentiation. Important aspects of meiotic development that distinguish it from mitotic growth include a highly increased rate of recombination, formation of the synaptonemal complex that aligns the homologous chromosomes, as well as separation of the homologues and sister chromatids during meiosis I and II without an intervening S-phase. Sporulation in yeast and spermatogenesis in higher eukaryotes are analogous developmental pathways that involve conserved genes, many of which are known to display transcriptional upregulation during the process. It has proved useful to compare the genome-wide expression patterns of model systems to identify and characterize genes required for sexual reproduction in all eukaryotes. This review will briefly outline various aspects of the microarray technology and then focus on its impact on our understanding of meiosis and gametogenesis in yeasts as well as multicellular organisms.

Microarrays and bioinformatics

High density oligonucleotide microarrays (GeneChips) and PCR microarrays are available for nearly all major model systems and *Homo sapiens* (Lockhart and Winzeler, 2000; Young, 2000). GeneChips contain oligonucleotide probes synthesized *in situ* on glass slides using photolithography and combinatorial chemistry (Lockhart *et al.*, 1996). These probes recognize transcripts from a single sample (Fig. 1a). PCR micro-

arrays are standard microscopic glass slides onto which PCR fragments or oligonucleotides are spotted (see <http://www.gene-chips.com>; Schena *et al.*, 1995; Lashkari *et al.*, 1997). The arrays are hybridized with a mixture of cRNAs from two samples, each labelled with a different fluorophor (CY3 and CY5; Fig. 1b). High density oligonucleotide microarrays yield a single fluorescence intensity measurement for each gene, whereas PCR microarrays produce a ratio between two signals (red and green) measured simultaneously in one spot. When all loci from a genome are represented on the array, it is possible to determine the transcriptome of a process, that is to say all transcripts present in a cell at a given time point are detected. The MIAME convention (Minimal Information About a Microarray Experiment) requires authors to provide raw data as well as information about their experiments (Brazma *et al.*, 2001) to standardize array data analysis. Normalized or scaled data are analysed with filtering algorithms to identify differentially regulated transcripts (Hastie *et al.*, 2000, 2001; Lockhart and Winzeler, 2000). Cluster algorithms can then be used to group expression patterns according to their similarity. This procedure is useful because loci of similar functions tend to be expressed in a similar way (Eisen *et al.*, 1998). Numerous cluster algorithms are available but none of them is perfectly adapted to all experimental conditions and approaches (Quackenbush, 2001; Table 1).

How reliable are microarray experiments? It is noteworthy that only a fraction of cell-cycle or meiotically regulated genes have been commonly identified by groups studying mitosis (Cho *et al.*, 1998; Spellman *et al.*, 1998) or sporulation (Chu *et al.*, 1998; Primig *et al.*, 2000), which may raise doubts about the overall

*Correspondence
Email: michael.primig@unibas.ch

Fig. 1. Expression-profiling experiments using high density oligonucleotide (GeneChip) and PCR microarrays. (a) Total RNA from samples to be compared (for example, treated versus untreated, as indicated) is prepared and reverse transcribed into the so-called target cRNA in the presence of nucleotide analogues coupled to biotin. The target is hybridized to the GeneChip, removed and the array is rehybridized with the streptavidine-phycoerythrin (SAPE) fluorophore that binds to biotinylated cRNA. Fluorescence intensities are computed from a scanned GeneChip image for each of 11–16 gene-specific probe pairs that contain perfect match (PM) and mismatch (MM) oligonucleotides (25 bases in length). MM oligonucleotides reveal non-specific hybridization signals because they contain a point mutation at position 13 that destabilizes the hybridization reaction (top and bottom row, see small insert). (b) cRNA samples are labelled with Cy3/Cy5 fluorophores, mixed and hybridized to one single array. A ratio between reference (untreated) and experiment (treated) is computed from one array image.

reproducibility of expression-profiling data. A possible explanation for poor overlap in the case of cell-cycle regulation is that Cho *et al.* (1998) relied on visual inspection of a single data set, whereas Spellman *et al.* (1998) used statistical data analysis of four time-course experiments, each using a different method to synchronize cells. It is possible that stress-based synchronization protocols are a source of error themselves (Cooper and Shedden, 2002). In the case of sporulation, both studies carried out thorough statistical and cluster analysis; therefore, the difference may be related to distinct array technologies and different sporulation efficiency (Table 2).

Whole genome expression-profiling meiosis and spore formation in yeasts

Saccharomyces cerevisiae has been extensively used for the analysis of mitosis and meiosis as well as gamete

formation (sporulation). Important genes were organized into five categories according to the timing of their induction: very early, early, middle, mid-late and late (for review, see Kupiec *et al.*, 1997). These results were complemented and extended by analysing the transcriptome of meiosis and spore development with GeneChips in two different genetic backgrounds including wild-type and meiosis-deficient strains. Nine hundred core genes were demonstrated to be meiotically regulated in a strain-independent manner, including hundreds that are also expressed during mitotic growth. The classical categories were further refined into seven expression clusters (Primig *et al.*, 2000). About 500 genes were found to be upregulated during sporulation in a similar analysis using PCR microarrays (Chu *et al.*, 1998). In both studies it emerged that timing of induction coincided in many cases with time of function, notably for genes involved in recombination (*SPO11*, *REC102*), pachytene checkpoint function (*PCH2*), sister chromatid separation

Table 1. URLs of important websites

Type and name	Uniform Resource Locator (URL)
Genomes	
GeneOntology	http://www.geneontology.org
NCBI	http://www.ncbi.nlm.nih.gov/Genomes/index.html
KEGG	http://www.kegg.org
GOLD	http://wit.integratedgenomics.com/GOLD
Bioinformatics	
The European Bioinformatics Institute	http://www.ebi.ac.uk
NCBI	http://www.ncbi.nlm.nih.gov
Swiss Institute of Bioinformatics	http://www.isb-sib.ch
Bioinformatics.org	http://bioinformatics.org
Microarrays	
Gene-Chips.com	http://www.gene-chips.com
GeneChip	http://www.affymetrix.com
Software	
Bioconductor	http://www.bioconductor.org/
ScanAlyze or Cluster	http://rana.lbl.gov/EisenSoftware.htm
CTWC server	http://ctwc.weizmann.ac.il
GeneSpring	http://www.sigenetics.com
GeneSight	http://www.biodiscovery.com
Spotfire	http://www.spotfire.com/
Array data repositories	
Gene Expression Omnibus	http://www.ncbi.nlm.nih.gov/geo
ArrayExpress	http://www.ebi.ac.uk/arrayexpress
DDBJ	http://www.ddbj.nig.ac.jp
Expression data	
Stanford Expression Connection	http://genome-www4.stanford.edu/cgi-in/SGD/expression/expressionConnection.pl
MIPS	http://mips.gsf.de/cgi-bin/proj/expression/start.pl
HuGE	http://www.hugeindex.org
Knowledgebases	
<i>Saccharomyces</i> Genome Database	http://genome-www.stanford.edu/Saccharomyces/
WormBase	http://www.wormbase.org/
FlyBase	http://www.flybase.org
Mouse Genome Informatics	http://www.informatics.jax.org
Gametogenesis databases	
GermOnline	http://germonline.unibas.ch
Ovarian Kaleidoscope	http://ovary.stanford.edu
TestisBank	http://medweb.uni-muenster.de/TestisBank/links.html
WormSperm	http://www.mcb.arizona.edu/wardlab/
dbZach Toxicogenomics Database	http://dbzach.fst.msu.edu
Microarray publication portals	
Stanford Microarray Database	http://genome-www5.stanford.edu/MicroArray/MDEV/publications.html
Affymetrix	http://www.affymetrix.com/community/publications/1.advanced_search.affx

(*REC8* or *SPO69*), regulation of meiosis I (*SPO13*), induction of middle meiotic genes (*NDT80*) and spore wall formation (*DIT1*).

A third attempt to determine the meiotic transcriptome of yeast using PCR microarrays, although it identified a subset of important core genes, failed to yield a large overlap with previous experiments (W. Lin and A. Nicolas, personal communication). It is therefore a distinct possibility that the actual core group of meiotically regulated genes may be in the range of only 150–200 loci. A potential way to tackle this problem is to

ask if there are genes the meiotic regulation or function of which is similar not only among array experiments but also among species. Clues to this critical question emerge from the transcriptome of sexual differentiation in *Schizosaccharomyces pombe* (Mata *et al.*, 2002). Fission yeast is only distantly related to budding yeast, which is why sequences and patterns conserved between these two organisms are considered significant in terms of evolution. In *S. pombe*, as in budding yeast, about 1000 genes are meiotically regulated. However, the group of similarly regulated meiotic genes is rather small,

Table 2. Summary of the species, biological focus and web addresses of microarray studies relevant for mitosis, meiosis and gametogenesis

Organism	Biological focus	URL	Reference
<i>Saccharomyces cerevisiae</i>	Mitotic cell cycle	http://171.65.26.52/yeast_cell_cycle/cell_cycle.html	Cho <i>et al.</i> (1998)
<i>S. cerevisiae</i>	Mitotic cell cycle	http://genome-www.stanford.edu/cell_cycle/	Spellman <i>et al.</i> (1998)
<i>S. cerevisiae</i>	Meiosis and sporulation	http://www.bioz.unibas.ch/primig/transcriptome	Primig <i>et al.</i> (2000)
<i>S. cerevisiae</i>	Meiosis and sporulation	http://cmgm.stanford.edu/pbrown/sporulation	Chu <i>et al.</i> (1998)
<i>Schizosaccharomyces pombe</i>	Meiosis and sporulation	http://www.sanger.ac.uk/PostGenomics/S_pombe/projects/sexualdifferentiation	Mata <i>et al.</i> (2002)
<i>Caenorhabditis elegans</i>	Gametogenesis	http://cmgm.stanford.edu/~kimlab/germline	Reinke <i>et al.</i> (2000)
<i>C. elegans</i>	Development	http://cmgm.stanford.edu/~kimlab/dev	Jiang <i>et al.</i> (2001)
<i>Drosophila melanogaster</i>	Testis expressed sequence tags	http://www.niddk.nih.gov/intram/people/boliver.htm#publications	Andrews <i>et al.</i> (2000)
<i>Mus musculus</i>	Development	http://expression.gnf.org	Su <i>et al.</i> (2002)
<i>M. musculus</i>	Development	http://read.gsc.riken.go.jp	Miki <i>et al.</i> (2001)
<i>M. musculus</i>	Testis arrays	http://genomebiology.com/2001/2/4/research/0014	Rockett <i>et al.</i> (2001)
<i>Homo sapiens</i>	Mitotic cell cycle	http://www.salk.edu/docs/labs/chipdata	Cho <i>et al.</i> (2001)
<i>H. sapiens</i>	Mitotic cell cycle	http://genome-www.stanford.edu/Human-CellCycle/Hela	Whitfield <i>et al.</i> (2002)
<i>H. sapiens</i>	Development	http://expression.gnf.org	Su <i>et al.</i> (2002)
<i>H. sapiens</i>	Spermatogenesis	http://compbio.med.wayne.edu/microarray	Ostermeier <i>et al.</i> (2002)

comprising only 75 genes, among which are regulators of cell-cycle progression (B-type cyclins like *CLB1* or *cdc13*, *CLB2-CLB6* or *cig2*), the seven components of the anaphase-promoting complex, as well as loci involved in recombination (*REC114* or *rec7*, *REC8* or *rec8*, *DMC1* or *dmc1*), sister chromatid cohesion (*SMC3* or *smc3*) and chromosome synapsis (*HOP2* or *meu13*). The overall process of meiosis and spore formation is similar in both fission and budding yeast but the regulatory machinery that orchestrates it is not. The corollary is that complex meiotic transcriptional regulation may have evolved at later stages to optimize the process and to meet the needs of multicellular organisms (Mata *et al.*, 2002). It must be emphasized that an important limitation of this comparative approach is the fact that *S. pombe* does not form the tripartite synaptonemal complex, a common feature of meiosis in most eukaryotes (Davis and Smith, 2001). It therefore lacks numerous structural components and enzymes involved in the establishment and maintenance of this important structure.

Meiotic regulation in budding yeast

To date, only three site-specific DNA-binding transcription factors (Ume6, Abf1 and Ndt80) involved in meiotic gene induction have been described (Gailus-Durner *et al.*, 1996; Chu and Herskowitz, 1998; Hepworth *et al.*, 1998). Ume6 was first identified in a screen for mutants that expressed meiosis-specific genes during mitotic growth (unscheduled meiotic expression) and was later demonstrated to encode a key regulator of nitrogen repression and meiosis (Bowdish *et al.*, 1995; Steber and Esposito, 1995). Ume6 exerts its function through direct interaction with a conserved

histone deacetylase complex that contains Rpd3, Sin3 and Iws2 or by binding to the meiotic activator Ime1 (Kadosh and Struhl, 1998; Goldmark *et al.*, 2000; Washburn and Esposito, 2001). An expression-profiling study of Ume6 deletion mutants compared with a wild-type strain revealed significant mitotic de-repression of almost 80 genes (Williams *et al.*, 2002). Among those whose functions are known, most participate in carbon and nitrogen metabolism or sporulation and approximately half of the genes directly regulated by Ume6 are induced during sporulation. Many genes are essential for various steps in the meiotic pathway including recombination (*DMC1*, *SPO11*), formation of the synaptonemal complex (*HOP1*, *HOP2*, *ZIP2*), sister chromatid cohesion (*SPO69* or *REC8*), the pachytene checkpoint (*PCH2*), control of meiosis I (*SPO13*) and prospore development (*MPC54*, *SMA1*).

A remarkable outcome of microarray experiments is that the 5'-upstream regions of several hundred meiotically upregulated loci lack matches to the target sites of Ume6, Abf1 and Ndt80 (Primig *et al.*, 2000). Moreover, numerous middle meiotic genes (clusters 5 and 6) are not regulated by Ndt80 (Chu *et al.*, 1998) and many early meiotic genes (cluster 4) are not mitotically de-repressed in Ume6 mutants (Williams *et al.*, 2002). This finding strongly indicates that meiotic expression involves as yet unknown regulators. To learn more about meiotic induction one needs to identify genes that display robust and highly correlated expression patterns in different strains. It is important to bear in mind that meiotic regulation as defined for 900 core genes includes those that increase in one strain and decrease in the other (Primig *et al.*, 2000). However, the fact that most of the known meiotic genes show highly correlated expression profiles indicates that

Fig. 2. A selection of reproducibly induced meiotic core genes. Three hundred and fifteen genes that display robust and strain-independent upregulation during meiosis and spore development in *MATa/α* wild-type but not *MATa/a* or *MATα/α* meiosis-deficient strains were retrieved from GermOnline (<http://germonline.unibas.ch>) on the basis of displaying a correlation coefficient ≥ 0.8 . The expression profiles as determined in SK1 *MATa/α* are used to group the genes together using the GeneTree function of GeneSpring 4.2.1. The Spearman Confidence algorithm was used with the separation ratio set at 0.5 and the minimum distance set at 0.001. Subsequently, the expression data obtained with the three other strains were visualized. Lanes, columns and sub-columns correspond to genes, strains and time points, respectively, as indicated. Vertical bars outline the cellular functions that represent the majority of the known genes falling into the expression clusters symbolized by grey boxes (C1–C7 are expression clusters 1–7). Stress: summary of functions involved in nutrient uptake as well as amino acid uptake and biosynthesis; DS: DNA synthesis; Rec: recombination; SC synaptonemal complex; MI + II: meiotic divisions; ascus: spore and ascus formation or maturation.

similar transcriptional patterns in different strains are likely to reflect biologically significant expression signals. A selection of 315 clustered core genes the microarray signals of which fulfil the criterion of reproducibility are shown (Fig. 2). The genes are upregulated only in sporulating and not in meiosis-deficient control strains and therefore their promoters are excellent candidates for further study. This group also includes 20 genes that show a clear pattern of cell-cycle regulation during mitotic growth; for example, *SPO12* which is essential for meiosis and required for mitotic growth in the absence of *DBF2*; *CLB5* and *CLB6*, which are important for pre-meiotic DNA synthesis; and *PDS1*, a regulator of cell-cycle progression (Klapholz and Esposito, 1980; Smith *et al.*, 2001; Agarwal and Cohen-Fix, 2002). The expression profiles of these genes during mitotic growth

and meiotic development are summarized (Fig. 3). It is not known how regulatory elements that mediate cell-cycle regulation are re-programmed into conferring a pattern of meiotic induction (and often re-repression). A better understanding of the promoters of these genes will provide insight into the molecular machinery that governs the transition from cell-cycle progression to meiotic differentiation.

Does meiotic expression imply meiotic function?

How many meiotically regulated genes are essential for the process? Whereas most important sporulation genes are meiotically upregulated (Kupiec *et al.*, 1997), it was demonstrated that genes that are meiosis-specifically expressed or induced are not necessarily essential for the

Fig. 3. Re-programming cell-cycle specific induction during meiotic development. The mitotic and meiotic expression patterns of 20 selected genes that display cell-cycle regulation are shown. Genes were clustered as described in Fig. 2 on the basis of the cell-cycle data from Cho *et al.*, 1998 (produced with GeneChips). The cell-cycle cluster was then used to visualize expression data obtained in sporulating SK1 and W303 strains as indicated. The expression profiles are shown for two cell cycles and meiotic development in two different strains. Cell-cycle-regulated genes that fall into the seven meiotic expression clusters (C1–C7) are colour coded.

process (Percival-Smith and Segall, 1986). Therefore, it was not surprising that only approximately 10% of about 300 upregulated genes displayed an essential phenotype in a large-scale deletion study (Rabitsch *et al.*, 2001). However, the correlation between meiotic regulation and essential function is statistically significant, hence expression profiling is valuable for developmentally important loci (Deutschbauer *et al.*, 2002). It is unclear why the majority of induced genes have no phenotype, but functional redundancy may provide an explanation in a number of cases (U. Schlecht, S. Brachat, P. Philippsen and M. Primig, unpublished). To date, approximately 400 loci are known to be required for meiosis and sporulation in yeast and this number will further increase as soon as the sporulation phenotypes of all non-essential yeast-deletion strains are available (Kupiec *et al.*, 1997; Briza *et al.*, 2002; Deutschbauer *et al.*, 2002; Giaever *et al.*, 2002; A. Nicolas, unpublished). Many of these genes are meiosis specific and

conserved and therefore could be candidates for important functions in higher eukaryotes. A selection of 75 yeast genes that are meiosis-specifically expressed and their potential homologues and known orthologues in fly, worm and mammals are summarized at <http://www.biozentrum.unibas.ch/personal/primig/gametes>.

Expression profiling of meiosis and the germline in higher eukaryotes

Identifying meiotic- and germ-cell-specific transcripts in multicellular organisms is a complex task because gonads contain different types of cell, only a fraction of which are germ cells. However, it is possible to obtain informative expression data using microarrays by comparing, for example, mutant worm strains devoid of male or female gametes with wild-type strains, by profiling gonadal tissues or by analysing purified germ cells at various stages of development (Table 2). The

following section of this review summarizes the outcome of recent experiments of these types using *Caenorhabditis elegans* and mammalian model systems.

The worm is an excellent organism for expression-profiling experiments because of the wealth of mutant strains available. Moreover, currently available microarrays cover nearly all annotated genes (Hill *et al.*, 2000; Jiang *et al.*, 2001; Reboul *et al.*, 2001). A landmark study compared a wild-type hermaphrodite to mutants devoid of any germ cells (*glp-4*) at three larval stages and one adult stage. In addition, the authors compared adult animals producing only spermatozoa (*fem-3(gf)*) with animals producing only oocytes (*fem-1(lf)*) (Reinke *et al.*, 2000). Using statistical analysis, the authors identified 1416 germline-expressed genes. Among those they found 650 sperm-specific or -enriched, 258 oocyte-specific or -enriched and 508 germline intrinsic loci that are expressed in both types of germ cells. An interesting outcome was that sperm-enriched loci contained a disproportionally high number of kinases and phosphatases as well as enzymes implicated in energy metabolism. This finding is in keeping with the fact that post-translational regulation plays an important role in spermatogenesis and that sperm motility requires a lot of energy. In a similar way, oocyte-enriched genes were found to include genes possibly involved in eggshell formation and DNA replication. Expression profiling facilitates identification of genes important for gametogenesis in yeast and large-scale gene 'knock-down' experiments by RNA interference indicate that this is also true for the worm (Piano *et al.*, 2002; Walhout *et al.*, 2002).

Somatic and gametogenic tissues are distinguishable by their associated expression profiles (Andrews *et al.*, 2000; Miki *et al.*, 2001; Su *et al.*, 2002). Such exploratory approaches help identify genes involved in developmental and metabolic pathways, for example, glycolysis. Clustering loci that play a role in this pathway separates them into forms that are expressed specifically in muscle, liver or kidney and testis (Miki *et al.*, 2001); forms that are expressed specifically in the testis may contribute to energy production during spermiogenesis (Welch *et al.*, 1992). Important insight into germ-cell-specific transcription comes from Camerini-Otero and co-workers who compared the testicular expression patterns of approximately 15 000 mouse cDNAs in wild-type and homozygous Spo11-deletion mice (N. Smirnova and D. Camerini-Otero, personal communication). Mutants fail to initiate meiotic recombination, which triggers apoptosis of spermatocytes and oocytes, thus causing both males and females to be sterile (Baudat *et al.*, 2000; Romanienko and Camerini-Otero, 2000). A clear pattern of deregulation in the absence of Spo11 was observed for approximately 100 genes in juvenile mice at 15 days after birth. Furthermore, in the testes of adult Spo11 deletion mice, the expression of 23% of the genes was altered, indicating substantial transcriptional deregulation in

the absence of double-strand break formation during mammalian meiosis. A recent report describes 153 TISP (transcript induced during spermiogenesis) genes the expression of which is upregulated during sperm maturation in mice, including 80 that are testis specific (Fujii T *et al.*, 2002). The authors found 82 previously uncharacterized sequences among the 153 loci that were organized into three different types based upon their expression in juvenile and adult tissue. This study, as well as data obtained with PCR microarrays that contain 900 testis genes (Rockett *et al.*, 2001), will be helpful for future expression-profiling experiments.

It would be interesting to determine the transcriptome of spermatogenesis using purified male germ cells. Sedimentation or elutriation techniques are time-consuming processes during which the expression of many genes may change. However, it is possible to distinguish rat Sertoli cells from dividing spermatogonia and differentiating spermatocytes as well as round spermatids on the basis of global expression patterns of about 800 genes filtered out among approximately 9000 loci and ESTs (expressed sequence tags) present on U34A GeneChips (Affymetrix). Moreover, a number of previously identified testis-specific or -enriched transcripts from genes like Tnp1 or Tnp2 (transition nuclear proteins), Prm2 (DNA binding protein), Sp10 (sperm-specific protein) and Sycp3 (SC component) are clearly detectable in purified germ-cell populations (U. Schlecht, P. Demougin, R. Koch, M. Bellis, I. Coiffec, B. Jégou and M. Primig, unpublished). An additional approach might be to use *in vitro* cell systems undergoing partial spermatogenesis, because it is possible to extract high-quality RNA rapidly from cultured cells (Jégou *et al.*, 2001). An important advance is the recent demonstration of *in vitro* differentiation of telomerase immortalized type A mouse spermatogonia (Feng *et al.*, 2002). Critical issues are germ-cell synchrony and the fact that only about 60% of the spermatogonia reach the haploid stage, as subtle transcriptional changes may be masked by cells that fail to transit the pathway (Feng *et al.*, 2002). Non-invasive techniques may also prove useful for microarray studies of spermatogenesis. Encouraging results were reported by Ostermeier *et al.* (2002) who analysed 30 000 cDNAs using spermatozoa purified from human testicular as well as ejaculate samples. The data indicate that among at least 2686 detected ESTs, transcripts from all stages of spermatogenesis are present. In addition, the authors argue that spermatozoa may provide functions involved in embryonic development rather than simply being a vehicle for the male complement of chromosomes.

Is meiosis-specific transcription likely to be conserved in all eukaryotes? Clues about conservation of expression patterns between closely related species come from a large-scale profiling analysis of tissues in humans and mice (Su *et al.*, 2002). About half of 799 putative orthologue pairs displayed strikingly similar patterns

of transcription in humans and mice at a correlation coefficient of at least 0.6 (with 0 being unrelated and 1 being identical). Among those, the authors found 104 conserved loci expressed in testis including known genes like *Sox5* and *Tekt2*. This finding underlines the usefulness of expression-profiling studies in mice and rats, the results of which are likely to be applicable to their human homologues. However, evolutionary distance is critical; *SPO11*, for example, is meiosis specific in budding yeast but only germ-cell-enriched in mammals, in which it is also weakly expressed in thymus and brain tissues (Klapholz *et al.*, 1985; Keeney *et al.*, 1999). Interestingly, 114 out of 155 putative *S. cerevisiae* orthologues of human cell-cycle-regulated genes display a cyclical expression pattern in yeast too (Whitfield *et al.*, 2002). The finding that many genes are conserved and cell cycles regulated in species as distant as yeast and humans supports the notion that comparative expression profiling is a reasonable means of identifying important genes in organisms not accessible to genetic and biochemical analysis. Meiotic profiling studies using distantly related fungi identify only a small set of < 100 commonly induced genes (Priming *et al.*, 2000; Mata *et al.*, 2002). It is possible that the number of loci that constitute the genetic core of gametogenesis in all eukaryotes may be in that range.

Knowledgebase development

Data management and organization of knowledge has become a key application of information technology in biological research. The global GeneOntology project provides a framework of rules and nomenclature for genes across species (Ashburner *et al.*, 2000). A gene product is described by the categories biological process, molecular function and cellular component. On the basis of this convention a number of species-specific databases from major model organisms are being constructed and maintained by curators (Table 1). Providing up-to-date information about tens of thousands of genes remains a serious challenge but one way of handling knowledge is to share the burden among many scientists. This requires a database that authors can access through the internet to publish and update their findings. The gametogenesis field is at the forefront of biological knowledge management that uses this type of approach. An organ-based database called *The Ovarian Kaleidoscope* (OKdb) integrates published information on individual ovarian genes (Ben Shlomo *et al.*, 2002), whereas *Mammalian Reproductive Genetics* (MRG) will contain genes involved in oogenesis and spermatogenesis in mammals (V. Cassen and B. Braun, unpublished). The *GermOnline* database will initially cover 11 model systems as well as *H. sapiens*, providing a platform for scientists studying meiosis and gamete

formation to directly contribute and update information. *GermOnline* is organized by an international board of curators to ensure the highest possible scientific quality of the content of the database. The information is meant to be provided by users through an online submission or curation system whose form uses free text and GeneOntology keywords (M. Primig and The Board of Curators, unpublished). A prototype of the database that enables researchers to make contributions and currently provides access to relevant microarray expression data from various species is available at <http://germonline.unibas.ch/>.

Conclusion

There is no doubt that global standardization of microarray data management and analysis, cheaper and better microarrays, as well as improved experimental protocols will further increase the impact of expression-profiling studies on biomedical research. This is a truly exciting phase in which biology, physics, biochemistry, mathematics and information technology blend into an array of novel approaches to understand sexual reproduction.

The authors thank A. Nicolas and D. Camerini-Otero for communication of unpublished results and D. Wolgemuth as well as the members of the laboratory for critical reading of the manuscript. U. Schlecht is supported by a grant from the Swiss Institute of Bioinformatics.

References

Key references are identified by asterisks.

- Agarwal R and Cohen-Fix O** (2002) Phosphorylation of the mitotic regulator Pds1/securin by Cdc28 is required for efficient nuclear localization of Esp1/separase *Genes and Development* **16** 1371–1382
- Andrews J, Bouffard GG, Cheadle C, Lu J, Becker KG and Oliver B** (2000) Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis *Genome Research* **10** 2030–2043
- Ashburner M, Ball CA, Blake JA *et al.*** (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium *Nature Genetics* **25** 25–29
- Baudat F, Manova K, Yuen JP, Jasin M and Keeney S** (2000) Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11 *Molecular Cell* **6** 989–998
- Ben Shlomo I, Vitt UA and Hsueh AJW** (2002) Perspective: the ovarian kaleidoscope database-II. Functional genomic analysis of an organ-specific database *Endocrinology* **143** 2041–2044
- Bowditch KS, Yuan HE and Mitchell AP** (1995) Positive control of yeast meiotic genes by the negative regulator UME6 *Molecular and Cellular Biology* **15** 2955–2961
- Brazma A, Hingcamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J and Ansorge W** (2001) Minimum information about a microarray experiment (MIAME)-towards standards for microarray data *Nature Genetics* **29** 365–371
- Briza P, Bogengruber E, Thur A, Rutzler M, Munsterkötter M, Dawes I and Breitenbach M** (2002) Systematic analysis of sporulation phenotypes in 624 non-lethal homozygous deletion strains of *Saccharomyces cerevisiae*. *Yeast* **19** 403–422

- Cho RJ, Campbell MJ, Winzeler EA *et al.* (1998) A genome-wide transcriptional analysis of the mitotic cell cycle *Molecular Cell* **2** 65–73
- Cho RJ, Huang M, Campbell MJ *et al.* (2001) Transcriptional regulation and function during the human cell cycle *Nature Genetics* **27** 48–54
- Chu S and Herskowitz I (1998) Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80 *Molecular Cell* **1** 685–696
- *Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO and Herskowitz I (1998) The transcriptional program of sporulation in budding yeast *Science* **282** 699–705
- Cooper S and Shedden K (2002) Analysis of cell-cycle gene expression in *Saccharomyces cerevisiae* using microarrays and multiple synchronization methods *Nucleic Acids Research* **30** 2920–2929
- Davis L and Smith G (2001) Meiotic recombination and chromosome segregation in *Schizosaccharomyces pombe*. *Proceedings National Academy of Sciences USA* **98** 8395–8402
- *Deutschbauer A, Williams R, Chu AM and Davis R (2002) Parallel phenotypic analysis of sporulation and postgermination growth in *Saccharomyces cerevisiae*. *Proceedings National Academy of Sciences USA* **99** 15 530–15 535
- Eisen M, Spellman P, Brown P and Botstein D (1998) Cluster analysis and display of genome-wide expression patterns *Proceedings National Academy of Sciences USA* **95** 14 863–14 868
- Feng LX, Chen Y, Dettin L, Pera RAR, Herr JC, Goldberg E and Dym M (2002) Generation and *in vitro* differentiation of a spermatogonial cell line *Science* **297** 392–395
- Fujii T, Tamura K, Masai K, Tanaka H, Nishimune Y and Nojima H (2002) Use of stepwise subtraction to comprehensively isolate mouse genes whose transcription is up-regulated during spermiogenesis *EMBO Reports* **3** 367–372
- Gailus-Durner V, Xie J, Chintamaneni C and Vershon AK (1996) Participation of the yeast activator Abf1 in meiosis-specific expression of the HOP1 gene *Molecular and Cellular Biology* **16** 2777–2786
- Giaever G, Chu AM, Ni L *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome *Nature* **418** 387–391
- Goldmark J, Fazio T, Estep P, Church G and Tsukiyama T (2000) The lws2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p *Cell* **103** 423–433
- Hastie T, Tibshirani R, Eisen MB, Alizadeh A, Levy R, Staudt L, Chan WC, Botstein D and Brown P (2000) 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns *Genome Biology* **1** Research 0003
- Hastie T, Tibshirani R, Botstein D and Brown P (2001) Supervised harvesting of expression trees *Genome Biology* **2** Research 0003
- Hepworth SR, Friesen H and Segall J (1998) NDT80 and the meiotic recombination checkpoint regulate expression of middle sporulation-specific genes in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **18** 5750–5761
- *Hill A, Hunter C, Tsung B, Tucker-Kellogg G and Brown E (2000) Genomic analysis of gene expression in *C. elegans*. *Science* **290** 809–812
- Jegou B, Le Magueresse, Battistoni B and Gerard N (2001) Is the *in vitro* maturation of germ cells accelerated in co-culture with Sertoli cells? *Molecular and Cellular Endocrinology* **183** 195, 197
- *Jiang M, Ryu J, Kiraly M, Duke K, Reinke V and Kim SK (2001) Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proceedings National Academy of Sciences USA* **98** 218–223
- Kadosh D and Struhl K (1998) Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin *in vivo*. *Molecular and Cellular Biology* **18** 5121–5127
- Keeney S, Baudat F, Angeles M, Zhou ZH, Copeland NG, Jenkins NA, Manova K and Jasin M (1999) A mouse homolog of the *Saccharomyces cerevisiae* meiotic recombination DNA transesterase Spo11p *Genomics* **61** 170–182
- Klapholz S and Esposito RE (1980) Isolation of SPO12-1 and SPO13-1 from a natural variant of yeast that undergoes a single meiotic division *Genetics* **96** 567–588
- Klapholz S, Wadell CS and Esposito RE (1985) The role of the SPO11 gene in meiotic recombination in yeast *Genetics* **110** 187–216
- Kupiec M, Byers B, Esposito R and Mitchell A (1997) Meiosis and sporulation in *Saccharomyces cerevisiae*. In *The Molecular and Cellular Biology of the Yeast Saccharomyces* pp 889–1036 Eds EW Jones, JR Pringle and JR Broach. Cold Spring Harbour Laboratory Press, New York
- Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO and Davis RW (1997) Yeast microarrays for genome wide parallel genetic and gene expression analysis *Proceedings National Academy of Sciences USA* **94** 13 057–13 062
- Lockhart DJ and Winzeler EA (2000) Genomics, gene expression and DNA arrays *Nature* **405** 827–836
- Lockhart DJ, Dong H, Byrne MC *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays *Nature Biotechnology* **14** 1675–1680
- *Mata J, Lyne R, Burns G and Bähler J (2002) The transcriptional program of meiosis and sporulation in fission yeast *Nature Genetics* **32** 143–147
- *Miki R, Kadota K, Bono H *et al.* (2001) Delineating developmental and metabolic pathways *in vivo* by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays *Proceedings National Academy of Sciences USA* **98** 2199–2204
- Ostermeier G, Dix D, Miller D, Khatri P and Krawetz S (2002) Spermatozoal RNA profiles of normal fertile men *Lancet* **360** 772–777
- Percival-Smith A and Segall J (1986) Characterization and mutational analysis of a cluster of three genes expressed preferentially during sporulation of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **6** 2443–2451
- *Piano F, Schetter A, Morton D, Gunsalus K, Reinke V, Kim S and Kempthues K (2002) Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Current Biology* **12** 1959–1964
- *Primig M, Williams R, Winzeler E, Tevzadze G, Conway A, Hwang S, Davis R and Esposito R (2000) The core meiotic transcriptome in budding yeasts *Nature Genetics* **26** 415–424
- Quackenbush J (2001) Computational analysis of microarray data *Nature Reviews in Genetics* **2** 418–427
- Rabitsch K, Toth A, Galova M *et al.* (2001) A screen for genes required for meiosis and spore formation based on whole genome expression *Current Biology* **11** 1001–1009
- Reboul J, Vaglio P, Tzellas N *et al.* (2001) Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. *Nature Genetics* **27** 332–336
- *Reinke V, Smith H, Nance J *et al.* (2000) A global profile of germline gene expression in *C. elegans*. *Molecular Cell* **6** 605–616
- Rockett JC, Christopher Luft J, Brian Garges J, Krawetz SA, Hughes MR, Hee Kinn K, Oudes AJ and Dix DJ (2001) Development of a 950-gene DNA array for examining gene expression patterns in mouse testis *Genome Biology* **2** Research 0014
- Romanienko P and Camerini-Otero R (2000) The mouse Spo11 gene is required for meiotic chromosome synapsis *Molecular Cell* **6** 975–987
- Schena M, Shalon D, Davis RW and Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray *Science* **270** 467–470
- Smith KN, Penkner A, Ohta K, Klein F and Nicolas A (2001) B-type cyclins CLB5 and CLB6 control the initiation of recombination and synaptonemal complex formation in yeast meiosis *Current Biology* **11** 88–97
- Spellman P, Sherlock G, Zhang M, Iyer V, Anders K, Eisen M, Brown P, Botstein D and Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of *Saccharomyces cerevisiae* by microarray hybridization *Molecular Biology of the Cell* **9** 3273–3297
- Steber CM and Esposito RE (1995) UME6 is a central component of a developmental regulatory switch controlling meiosis-specific gene expression *Proceedings National Academy of Sciences USA* **92** 12 490–12 494
- *Su A, Cooke M, Ching K *et al.* (2002) Large-scale analysis of the human

- and mouse transcriptomes *Proceedings National Academy of Sciences USA* **99** 4465–4470
- *Walhout A, Reboul J, Shtanko O *et al.* (2002) Integrating interactome, phenome, and transcriptome mapping for the *C. elegans* germline *Current Biology* **12** 1952–1958
- Washburn BK and Esposito RE (2001) Identification of the Sin3-binding site in Ume6 defines a two-step process for conversion of Ume6 from a transcriptional repressor to an activator in yeast *Molecular and Cellular Biology* **21** 2057–2069
- Welch JE, Schatte EC, O'Brien DA and Eddy EM (1992) Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse spermatogenic cells *Biology of Reproduction* **46** 869–878
- *Whitfield M, Sherlock G, Saldanha A *et al.* (2002) Identification of genes periodically expressed in the human cell cycle and their expression in tumors *Molecular Biology of the Cell* **13** 1977–2000
- *Williams R, Primig M, Washburn B, Winzeler E, Bellis M, Sarrauste de Menthère C, Davis R and Esposito R (2002) The Ume6 regulon coordinates metabolic and meiotic gene expression in yeast *Proceedings National Academy of Sciences USA* **99** 13 431–13 436
- Young R (2000) Biomedical discovery with gene arrays *Cell* **102** 9–15