Review TRENDS in Parasitology Vol.20 No.6 June 2004

Microarrays and stage conversion in *Toxoplasma gondii*

John Ellis[1], Duncan Sinclair[1] and David Morrison[2]

[1]Department of Cell and Molecular Biology, University of Technology Sydney, Westbourne St, Gore Hill, Sydney, NSW 2065, Australia
[2]Department of Parasitology (SWEPAR), National Veterinary Institute and Swedish University of Agricultural Sciences, SE-751 89 Uppsala, Sweden

Microarray technology is emerging as a powerful method for the study of those changes in gene expression that occur when a parasite converts from one life cycle stage to another. The use of appropriate experimental design and data analysis methodologies is essential if this technology is to be correctly exploited. In this review, the study of stage conversion in the transformation of *Toxoplasma gondii* tachyzoites into bradyzoites is considered as an example of where the application of microarrays is revealing exciting new knowledge on underlying molecular mechanisms. The more general application of microarrays in parasitology is also discussed.

Two closely related, cyst-forming coccidian parasites, *Toxoplasma gondii* and *Neospora caninum*, are economically important pathogens of livestock: *T. gondii* causes fetal loss in sheep and *N. caninum* is an abortifacent in cattle. *Toxoplasma gondii* is also known to occur abundantly in humans, in whom it can cause life-threatening disease in the fetus and in immunosuppressed individuals. The life cycles of the two taxa are similar, involving alternating modes of infection of the intermediate host (such as sheep or cows) or the definitive host (cats for *T. gondii* and dogs for *N. caninum*). However, one of the most distinguishing features of their life cycles is their persistence in tissues of an intermediate host as cystic stages, which is thus one of the reasons for their classification in the Sarcocystidae.

Tissue cysts are typically found in the central nervous system (CNS) or musculature, where they can persist for long periods of time, leading to long-term infection. Reactivation of cystic stages of *T. gondii* in an immunocompromised individual can result in clinical disease such as cerebral encephalitis [1], whereas recrudescence of *N. caninum* is thought to be associated with cases of sporadic abortion in cattle [2]. Understanding the mechanisms leading to cyst formation (cystogenesis) and their reactivation is therefore conceptually important.

Stage conversion and microarrays

Cyst formation requires the transformation of the rapidly growing tachyzoite form of the parasite to the comparatively dormant bradyzoite form. Such stage conversion in *T. gondii* has been studied for some years using either *in vivo*-generated cysts or *in vitro*-generated bradyzoites. Tachyzoites can be easily cultured through infection of a host cell and will differentiate into bradyzoites, and eventually cysts, under a variety of conditions *in vitro* [e.g. high culture pH [2,3], heat shock [2], intracellular nitric oxide (NO) [4,5], respiratory pathway inhibitors or chemicals [4,6]. *In vivo*, cyst formation occurs in response to interferon (IFN)-γ-mediated immunity [7]. The central dogma that has evolved from these studies is that a reduction in growth rate of the parasite results (later or simultaneously) in a switch from tachyzoite to bradyzoite, and hence the promotion of cystogenesis. This process (Figure 1) is associated with the disappearance of tachyzoite-associated markers (such as SAG1) [8,9] and the appearance of bradyzoite and cyst-related markers such as BAG1, MAG1, SAG4 and CST1 [10,11]. A great deal has been written on this specific topic [11–14], which will not be reiterated here; however, it is sufficient to point out that the temporal expression of these bradyzoite-expressed genes appears similar in all *in vitro* methods used to investigate stage conversion at the molecular level.

Microarrays are being extensively used now in parasitology to profile gene expression of taxa on a genome-wide scale (Box 1). The general principles behind the use of microarrays, and their strengths and weaknesses, have been described in these pages previously [15], although

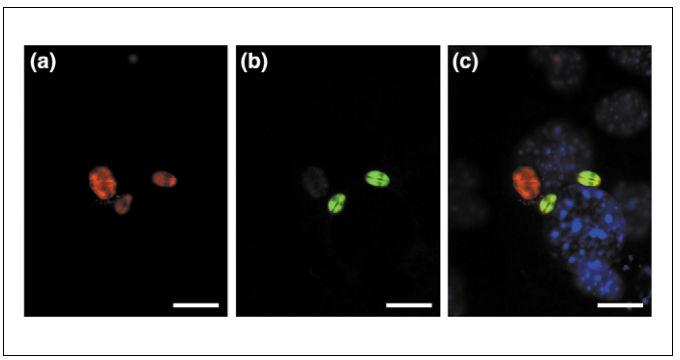

Figure 1. *In vitro* stage conversion from tachyzoites to bradyzoites. Host cells (such as Vero) are infected with tachyzoites and cultured under conditions that promote stage conversion into bradyzoites and intracellular cysts. The transformation process is followed using a double-immunofluorescent staining method with antibodies to: **(a)** SAG1 (stained red); or **(b)** BAG1 (stained green). In **(c)**, the specimen is stained with both antibodies and viewed under ultraviolet light. This figure was kindly provided by A. Hemphill and N. Vonlaufen (Berne). Scale (a–c) = 20 μm.

Corresponding author: John Ellis (j.ellis@uts.edu.au).

 1471-4922/$ - see front matter doi:10.1016/j.pt.2004.04.006

Box 1. Microarrays and their applications

The measurement of gene expression on a genomic scale has been made possible with the advent of array technology [52,53]. Arrays currently take two forms: filter-based gene arrays and gene chip technology. Filter-based arrays are commercially available and consist of charged membranes containing hundreds of duplicate spots of well-characterized cDNAs (e.g. the Atlas expression arrays distributed by Clontech). Microarrays are chips, the size of microscope slides, on which are spotted arrays of cDNA fragments or oligonucleotides. These 'spots' on the array represent the sequences of specific genes from the organism under study. Figure I shows the typical steps that occur during microarray analysis. Reverse transcriptase (RT)-PCR is used to synthesize cDNA from mRNA, and the cDNA is then hybridized to the nucleotides on the array. The extent of hybridization (determined by measurement of spot fluorescence using a microarray scanner) reflects the degree of expression of that gene in the mRNA population. Microarrays enable comparison of the relative levels of gene expression in two samples that are hybridized simultaneously to the same array. In this way, the upregulation or downregulation of gene expression in the experimental sample (e.g. derived from *in vitro*-generated bradyzoites) relative to the control (from tachyzoites) is determined by the relative binding of the two differentially labeled cDNAs to the same spot in the array.

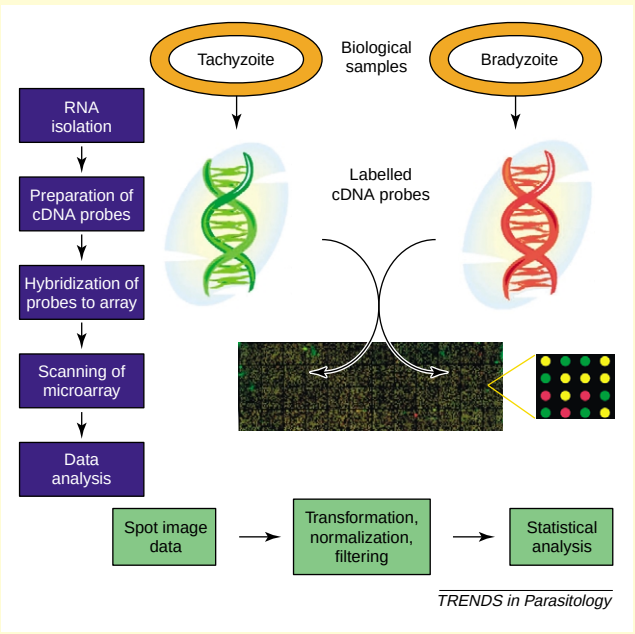

Figure I. Steps involved in the analysis of gene expression using microarrays.

Although the reported uses of microarrays are varied and numerous, they are being used increasingly in the study of infectious diseases; for example, in the study of host gene expression in response to infection by bacteria, viruses or parasites. Gail *et al.* [54] used filter-based arrays to investigate the transcriptional response of human fibroblasts to *Toxoplasma gondii* infection. Out of 597 genes, they demonstrated increased expression of ~40 genes in response to infection, including the transferrin receptor and MacMARCKS mRNAs, whose response was specific to *T. gondii* infection. Only one gene, the CCAAT-binding transcription factor, was found to be downregulated. Similarly, Blader *et al.* [55] used microarrays containing cDNAs from ~22 000 human genes to investigate the response of human fibroblasts to *T. gondii* infection. They showed that, soon after infection (~2 h), an increase in transcription occurred, and many of the genes whose expression was upregulated are normally associated with the immune response. At 24 h post-infection, host glycolytic and mevalonate pathway genes were upregulated along with others, which is in keeping with the observation that *T. gondii* must scavenge glucose and cholesterol from its host.

In addition, microarrays are being used to study the many elements of the immune response to pathogens [56,57]. For example, it was demonstrated using microarrays that dendritic cells responded transcriptionally to three different microbes with a core set of response genes plus a pathogen-specific set of genes [58]. Similarly, in response to six different bacterial stimulants, human peripheral blood mononuclear cells showed a common induction of a variety of cell–cell signaling, immunoregulatory and proinflammatory mediators [59]. Of particular importance here is the observation that microarrays could be used to survey and compare the response of the immune system to different forms of the pathogen (e.g. killed and live organisms). Indeed, live organisms were shown to modulate actively the host response when compared with killed bacteria [59].

Custom microarrays have been developed for *Schistosoma mansoni* and used to identify gender-associated gene transcripts, thereby enhancing our understanding of *Schistosoma* sexual biology [60]. *Trypanosoma brucei* microarrays [61] were used for identification of genes differentially expressed according to life cycle stage, and specific *Plasmodium falciparum* arrays [30,62,63] enabled the study of malarial parasite gene expression during the development of different life cycle stages. In the later studies, about half of the genes changed in expression levels during development, thereby illustrating the complexity of the processes that are involved in the rearrangement of the molecular composition of parasitic organisms. Surprisingly, even genes such as actin, which are normally considered as constitutively expressed, changed in their expression levels.

interested readers are also encouraged to look at introductory texts [16,17]. Cummings and Relman [18] provided one of the first, intriguing insights into how microarrays can be used to study microbes such as parasites, and the interactions between them and their hosts. In the first instance, microarray technology can be used simply to investigate changes in gene expression between two different cell populations, such as occurs during the conversion of tachyzoites to bradyzoites. However, it is also quite useful to try and identify groups of genes that have similar patterns of expression. Determining when and where genes are expressed allows one to make inferences about their function. For example, it has been recognized for some time that induction of bradyzoite formation *in vitro* results in the coordinated expression of bradyzoite-specific markers, along with the associated loss of tachyzoite-specific markers [10]. These observations have led, in part, to the raising of hypotheses that genes expressed early in the stage conversion will affect the differentiation pathway, whereas genes expressed later in the process might only affect bradyzoite maintenance and persistence. The application of microarray technology for investigating bradyzoite formation seeks to elucidate the identity and function of these intracellular processes.

Arguably, one of the greatest resources developed in parasitology over the past 10 years has been the establishment of databases containing sequence data derived from cDNAs of numerous parasites and their different life cycle stages. Large-scale sequencing of tachyzoite and bradyzoite expressed sequence tags (ESTs) have also been carried out for *T. gondii* [19,20] (and to a lesser extent for tachyzoites from *N. caninum* [21,22]), and these cDNA collections have now facilitated the production of

microarrays that are being used for profiling changes in gene expression during stage conversion of tachyzoites to bradyzoites. Database searching and clustering of the sequence data has allowed redundancy to be limited in the microarray design, thereby maximizing the number of genes that can be included in the experimental design. A second generation of microarrays based on gene-specific oligonucleotides will undoubtedly appear.

The hope of current research, based on significant levels of investment, is clearly to understand the molecular events that occur during stage conversion and, in the longer term, this might facilitate the production of technologies that can prevent this conversion from occurring. The development of a rapid method for the generation of tachyzoites from *T. gondii* or *N. caninum* that do not encyst would clearly be of significant commercial interest. A live vaccine (using the novel S48 strain) against *Toxoplasma*-associated abortion in sheep is available in selected countries of the world, showing the feasibility of marketing such a product that does not persist in animal tissues [23].

Experimental design and data analysis

Typically in microarray analyses, one is trying to identify genes that have statistically significant changes in the level of their expression under the experimental conditions employed. Morrison and Ellis [24] provide a detailed discussion of the experimental design and data analysis techniques currently employed for this purpose (Box 2). Genes are identified that might be over- or under-expressed in the RNA populations under study, and a twofold change in expression relative to a nominated reference sample is becoming entrenched in the literature. As discussed in Ref. [24], this approach is completely arbitrary and there is no theoretical basis for choosing any 'fold' change as standard. Therefore, appropriate quantitative analysis of the data should involve statistical hypothesis testing in addition to recognition of patterns by multivariate analysis. Unfortunately, there is no standard protocol for performing these analyses, and so it is appropriate to try and evaluate a variety of different data analysis techniques. Nevertheless, the aims of any microarray experiment design should be to ensure the generation of high-quality data by suitably preprocessing the raw data generated, and then testing various hypotheses by statistical or multivariate analyses.

The ability to identify rigorously genes or groups of genes that are interesting in terms of their patterns of gene expression is highly dependent on adequate replication being incorporated into the experimental design, to allow for all sources of variation that might obscure the patterns one is seeking to identify (Box 3). This could occur in the form of real biological variation between experimental treatments, or through some form of variation associated with the technology (e.g. extraction of RNA, or hybridization processes). It should be recognized that each array must be considered an independent experiment, and these are the true replicates for statistical purposes. Consequently, adequate numbers of arrays should be included in the microarray experiment, preferably in a design that is statistically efficient. It is the demonstration of a common

Box 2. Microarray data analysis

Analysis of microarray data can be directed towards three distinct purposes: class comparison, class prediction or class discovery. In class comparison, data are used to test explicit hypotheses about differences between experimental groups of samples; for example, one might test whether some experimental treatment has an effect on gene expression compared with a control treatment. In class prediction, the objective is to use the data to predict the characteristics of unknown specimens based on the characteristics of known specimens; for example, an unknown DNA sequence can be assigned to one of a set of known gene classes. In class discovery, general mathematical patterns are investigated in the data, which might then be interpreted as having some biological meaning; for example, it might be possible to distinguish between different functional classes of genes based on their expression patterns.

For class prediction, a range of supervised multivariate data analysis techniques are available, and for class discovery an even broader range of unsupervised multivariate techniques have been developed. Each of these multivariate techniques has its own limitations, which need to be kept in mind when making a choice from among them. In particular, each technique searches for a different type of pattern in the data, and it might fail to find alternative but equally valid patterns. This is important because, in class prediction and discovery, the data patterns are not being specified beforehand, and it is thus necessary to choose techniques that are capable of finding the pattern(s) residing in each unique dataset. It is therefore a worthwhile strategy to examine the results of several different types of analyses, rather than merely choosing one and sticking to it – consistency of the output from a range of analyses is a good indicator that it represents a real biological pattern, whereas inconsistency is often indicative of mathematical artefacts from the particular analyses chosen. The choice of appropriate multivariate analysis techniques for microarray data is thus an area of active research.

For class comparison, several experimental groups are being directly compared. The best approach to this problem is to use traditional techniques such as analysis of variance. However, issues related to multiple hypothesis testing and probability estimation still need to be evaluated. In particular, a specific problem arises with the use of statistical tests for microarray data because a large series of statistical hypotheses are being tested (i.e. one for each of the thousands of genes) and, if a large number of statistical comparisons are made, then some of the decisions to reject the null hypothesis will be due to chance errors (false positives). Furthermore, the usual parametric assumptions that the data have a normal frequency distribution with equal variances are likely to be invalid for microarray data (i.e. the data will violate the assumptions of the analysis), and so these possibilities should not be used. Devising methods for dealing with these problems is also an active area of research.

Perhaps the most unusual aspect of data analysis for arrays is the need for normalization of the data before the analysis. Normalization is related to the technical nonreproducibility of array technology – the current manufacturing and operating characteristics are not yet good enough for unwanted sources of variation to be dealt with automatically in the laboratory protocol, and so the biologist needs to adjust the data afterwards to remove technical artefacts. Unfortunately, we do not yet know enough details about the possible causes of the artefacts for there to be a uniform recommended protocol that is known to be biologically appropriate. So, most of the suggestions are heuristic mathematical techniques (e.g. globalization, smoothing, variance stabilization) that seem to work but might have as-yet-unknown less-desirable qualities. Interestingly, all of the experiments discussed here used vector-derived sequences for normalization, which is quite different from the more mathematical approaches usually adopted.

Box 3. Replication of arrays

Irrespective of the form of data analysis, adequate replication of the arrays themselves is probably the most important current issue concerning the quantitative analysis of microarray data. Convincing evidence that a mathematical pattern is real, and not generated by random chance, can only be provided by consistency of gene expression across a series of independent experimental replicates, and the more variable the data are then the more replicates that are needed. Most microarray experiments performed to date have had inadequate replication to answer properly the experimental question being examined, and this point has been re-iterated in several recent reviews.

For illustrative purposes, Ref. [24] used analysis of variance to analyze some of the data of Ref. [27], as described in the main text, and this analysis can be used to examine quantitatively how much effect each of the various components of the experiment has on the observed variation in gene expression levels. The most important of the variance components for the four-factor mixed analysis of variance are shown in part (a) of Table I below (there are four other components, each of which contributes $<0.1\%$). This analysis indicates that most of the differences in observed gene expression levels are related to the different conditions under which the cultures were grown, which is one of the two experimental manipulations. However, of more importance for our purposes here, the variation owing to differences between the replicate arrays is unduly large, larger even than the differences between the two parasite strains being studied (which is the other experimental manipulation).

From an experimentalist's point of view, this is unacceptable because the arrays in this case are merely technical replicates, and this therefore indicates poor repeatability of the experimental manipulations. If we look at this problem in more detail, as shown in part (b) of the table, it is apparent that the experimental combination where the mutant parasite strain was grown under bradyzoite growth conditions is the culprit – the two replicate arrays for this part of the experiment are clearly not duplicates of each other. This situation is particularly unfortunate in this example, because this was the experimental combination that was predicted to have different gene expression levels compared with the other experimental combinations. The message is thus clear: array technology is currently too imprecise to be used with only a small number of technical replicates, in addition to whatever biological replicates are needed. It is therefore also important that the changes in gene expression observed in microarray experiments be confirmed by alternative methods such as reverse transcriptase (RT)-PCR.

Table 1. An example of how variation in gene expression among replicate arrays can detrimentally affect the data analysis[a]

Component of variation	Contribution to total variation (%)[b]
(a) Full analysis	
Strains (parental versus mutant)	4.9
Conditions (tachyzoite versus bradyzoite)	84.1
Expressed sequence tags (ESTs; 2066 different ESTs examined)	0.8
Strains * conditions interaction	2.8
Arrays (two replicates)	7.2
(b) Array component only	
Parental + tachyzoite	8.7
Parental + bradyzoite	20.1
Mutant + tachyzoite	2.8
Mutant + bradyzoite	68.3

[a]This example is based on the data set from Ref. [24].
[b]The data represent the relative percentage of how much of the total amount of variation in gene expression is 'caused' by each of the experimental components listed. This shows that $>7\%$ of the variation is due simply to differences between the duplicate arrays, and that $>68\%$ of this 7% is due to the nonreproducibility of the arrays treated with the sample from the mutant strain grown under bradyzoite conditions.

pattern in the gene expression across a series of replicates that will provide convincing evidence that the pattern is real and not generated by random chance. It is probable that a twofold change in gene expression will not be detected by the use of a single array, and that five arrays might be sufficient to detect a 1.5-fold change in expression. Power analyses can be used prospectively to assist in experimental design. Given knowledge on the magnitude of the likely changes in gene expression, power analysis can be used to predict the number of arrays necessary to detect the pattern expected.

Microarrays for studying stage conversion

Cleary *et al.* [25] reported a study on tachyzoite to bradyzoite stage conversion using microarrays constructed from ~4400 *T. gondii* bradyzoite cDNAs, representing at least 613 separate contigs; 95 abundant tachyzoite ESTs were included. Stage conversion from tachyzoite to bradyzoite of the BSG-4 clone of the Pru strain [26] of *T. gondii* was induced *in vitro* in human foreskin fibroblasts and changes in gene expression were investigated by comparing mRNA levels of *in vitro*-grown tachyzoites with mRNA levels of parallel cultures grown in high-pH media at 2, 3 and 4 days after bradyzoite induction. The expression patterns shown by the ESTs during bradyzoite induction were classified as induced, repressed, constitutive or intermediate according to the mean bradyzoite to tachyzoite transcript level determined for the EST or its contig represented on the microarray. Quantitative analyses showed that there was an inverse correlation between mRNA abundance in tachyzoites and bradyzoites for many developmentally regulated genes [25]. Independent confirmation of these relative changes in gene expression was undertaken by northern blotting. Of importance was the observation that genes previously recognized as being developmentally regulated, such as *SAG1*, *NTP1*, *BAG1*, *SAG4A* and *LDH2* were all identified by microarray analyses as being differentially expressed, thereby confirming that microarray technology could faithfully display the changes in gene expression that occur during stage conversion. Most of these genes encode either surface molecules or secretory molecules from organelles associated with invasion of host cells (*ROP, MIC*) or intracellular development (*GRA*). A subset

of the genes comprised novel, differentially expressed genes that could be identified by database searching and that were not previously recognized as being developmentally regulated. The remaining contigs had no significant homology to known genes.

Cluster analysis identified several distinct classes of regulated genes that differed in their expression patterns. Many genes were not just simply switched on or off during stage conversion. For example, genes that were induced in bradyzoites fell into five categories: (i) those with transient expression (Day 2 only; Days 2 and 3; or Day 3 only); (ii) those with expression in mid-to-late stages of the conversion process (Days 3 and 4); (iii) those with expression in late stages of the conversion process (Day 4 only); (iv) those with expression at all stages (Days 2, 3 and 4); and (v) those that were all highly expressed throughout stage conversion (Days 2, 3 and 4). It was suggested that those genes that were expressed early in the stage conversion process, such as *SAG2C/D* and *SAG4A*, could be essential drivers of bradyzoite formation.

In a related study using bradyzoite-specific expression of green fluorescence protein (GFP) as a selection strategy [26], mutant *T. gondii* was generated by chemical mutagenesis with ethylnitrosylurea and mutants (tachyzoite to bradyzoite differentiation mutants or Tbd^{-}) were identified by their reduced ability to express GFP three days after transfer to high-pH medium. These mutants, under conditions that promote stage conversion into bradyzoites, did not routinely express the CST1 cyst wall protein and had substantially reduced expression of well-recognized bradyzoite markers such as BAG1 and SAG4A. It was therefore concluded that these mutants had a global defect in the differentiation pathway leading to stage conversion. Using cDNA microarrays, the gene expression profiles of these mutants were investigated under conditions of stage conversion *in vitro*. Four different mutants were studied, each of which showed different characteristics in terms of their gene expression profiles. All mutants showed a decreased expression of those genes identified by Cleary *et al.* [25] as being induced during bradyzoite formation. Indeed, 83% of the genes that had decreased expression in the Tbd^{-} mutants are normally induced in the bradyzoite. Other categories of genes (e.g. enolase) were identified that were under-expressed in one or more mutant populations.

Other contigs were identified that are not normally induced during bradyzoite formation but were characterized by reduced expression in the Tbd^{-} mutants. These contigs were identified by their homology to 14-3-3, MAG1, ATPase A subunit, PITSLRE kinase, histone H2A variant, cyclophilin, ubiquitin-like protein and several ribosomal genes. It was hypothesized that such genes might have a regulatory role in bradyzoite formation, or might simply be dependent on common transcription factors.

Matrajt *et al.* [27] used insertional mutagenesis to generate stable populations of mutant parasites. A stage-specific positive/negative selectable marker based on a strong bradyzoite-specific promoter driving expression of the *T. gondii* hypoxanthine-xanthine-guaninephosphoribosyltransferase was used to transfect tachyzoites, and mutants in the bradyzoite differentiation pathway were identified by induction of bradyzoite formation *in vitro* and selection in 6-thioxanthine. One of the key advantages behind this approach is that insertional mutagenesis carried out in this way greatly facilitates the subsequent cloning and identification of the gene loci into which the insertion has occurred.

The mutant parasites grew normally even under conditions that would favor bradyzoite formation and, under culture conditions that favor stage conversion, they also showed reduced expression of markers normally associated with bradyzoites such as BAG1. Similar to the study of Singh *et al.* [26], RNA from a parental strain and a mutant (B7) strain grown under conditions of high pH were hybridized to bradyzoite cDNA microarrays to investigate changes in gene expression. The B7 mutant population grown under bradyzoite conditions was more similar in its patterns of gene expression to tachyzoites from either the parent or mutant strains.

Morrison and Ellis [24] illustrated the possibilities for microarray data analysis, along with flagging important experimental design issues [24]. The datasets from Refs [26] and [27] were analyzed for illustrative purposes. Although this study [24] concentrated on emphasizing the methodology of data analysis, it is interesting to make note of other important outcomes from this study. In the first instance, *T. gondii* mutants that failed to go through stage conversion expressed a variety of genes coding for tachyzoite-associated products such as NTPase, GRA1, GRA2, GRA5, GRA6, GRA7, undulin, toxfilin, SUL1 and ROP4. These mutants also had reduced levels of expression of the bradyzoite-specific antigens 18 and 65 kDa, SAG2- and SAG4-related mRNAs, LDH and a proteasome subunit. The main conclusion drawn from collective observations is that the four mutant strains described by Singh *et al.* [26] had a block in one or more regulatory pathways that prevented a commitment to bradyzoite formation from occurring, thereby giving a transcriptional profile that was 'tachyzoite-like'. Such a conclusion was first provided by Matrajt *et al.* [27], as this is the most parsimonious explanation for these observations.

Second, several of the *T. gondii* mutants were associated with repression of a wide variety of cDNA contigs coding for ribosomal proteins such as L18, L22, L24 and S27a. The role of the ribosome in stage conversion and regulation of gene expression through translational control has largely been overlooked, even though stage conversion clearly involves clear changes in cell shape and composition which, in turn, result from changes in RNA and protein synthesis. A survey of the literature on ribosomal proteins of parasitic organisms shows that many such proteins are differentially expressed during the life cycle of a parasite. For example, in *Eimeria tenella*, the mRNA coding for the S3a ribosomal protein is poorly expressed in the first-generation schizont, yet strongly upregulated in the second-generation merozoite [28]; the *Theileria annulata* S17 protein is differentially expressed between the macroschizont and piroplasm stages [29], and components of the translational machinery, including ribosome-associated genes, are also recognized as being

differentially expressed during development of *Plasmodium falciparum* blood stages [30]. Although such evidence suggests that translational control of gene expression might be important in parasitic organisms, harder evidence in support of this conclusion is needed for *T. gondii*. Yang and Parmley [31] demonstrated that lactate dehydrogenase (LDH) is developmentally regulated during stage conversion of *T. gondii*, and that expression of LDH1 (one of the two LDH isoforms) was under translational control as mRNA coding for it could be detected in the absence of the protein [31]. Translational control also occurs during the expression of the BRS4/P36 bradyzoite-specific surface antigen (discussed in Ref. [32]). Recently, it was shown that stage conversion was associated with a downregulation of translation initiation factor 4A expression, showing that bradyzoite formation also involves a reduction in translation initiation activity [33].

Some ribosomal proteins also possess a secondary function within a cell, in addition to their essential role as a ribosome component. Roles for ribosomal proteins in essential cellular processes such as cell proliferation, replication, transcription, RNA processing, DNA repair, regulation of translation, regulation of development and apoptosis have been reported [cited in Refs 28, 29], thereby indicating the complexity of functions contributed to by ribosomal proteins. Such functional studies have yet to be undertaken in parasitic organisms like *T. gondii*.

Finally, it was possible to identify genes that possessed similar patterns of gene expression in the mutants generated in the study described in Ref. [27]. For example, the expression of ESTs 3906, 4131 and LDH (4054) were similarly repressed in all four mutants examined. Thus, one of the big benefits to be gained from undertaking microarray analyses is that it can assist in gene discovery processes through identifying genes that possess similar patterns of gene expression under the conditions studied.

Stage conversion in *N. caninum*

The morphology (size and thickness of the cyst wall) of the tissue cyst stage of *N. caninum* and *T. gondii* are different, thereby allowing a provisional identification to be made, for example, during diagnostic investigations. Such phenotypic differences presumably reflect the different molecular compositions these stages have in the two taxa.

The ability of *N. caninum* tachyzoites to differentiate into bradyzoites *in vitro* has been recently reported [32,33]. Using the culture conditions established for *T. gondii*, cyst-like structures could be established, and reagents used to detect *T. gondii* cysts (such as antibodies to BAG1) were found to be crossreactive to these *N. caninum* cysts by immunofluorescence and western blotting [34,36]. The efficiency of cyst formation *in vitro* for *N. caninum* was not very high in comparison with *T. gondii*. An alternative *in vitro* culture method was described that uses continuous culture of *N. caninum*-infected murine epidermal keratinocytes for up to 8 days in the presence of 70 μM sodium nitroprusside (SNP) [35]. The expression of bradyzoite markers was confirmed by immunofluoresence, along with vacuolar staining with *Dolichos biflorans* agglutinin. SNP treatment also led to the downregulation of expression of the major tachyzoite surface antigen NcSAG1. A modified version of this protocol has been developed for Vero cells using lower concentrations of SNP, because it proved difficult to purify bradyzoites from keratinocytes [36].

The availability of sequence and microarray datasets in the public domain allows their analysis using a variety of methods, including the mining of databases for genes from other taxa that have not been previously recognized as important in stage conversion. Through analyses of *T. gondii* datasets, and subsequent mining of *N. caninum* sequence databases, it is possible to identify *N. caninum* genes that are implicated in stage conversion events in this species, thus allowing us to fast-track hypotheses on cyst formation in this taxon. It comes as no surprise that homologs of many well-known and well-characterized genes of *T. gondii* exist in *N. caninum* databases. BAG1, LDH and SAG4 are easily identifiable by database searching of databases, but only further investigations will tell if the molecular mechanisms of cyst formation are similar between the two taxa.

Signaling

Prior knowledge on the mechanism of tachyzoite to bradyzoite transformation and the formation of cysts had shown that the changes in cell shape and composition were associated predominantly with control of gene expression at the transcriptional level. The available evidence demonstrated, through detection of changes in mRNA level by northern blotting and/or reverse transcriptase (RT)-PCR, that stage conversion involved the up- or downregulation of gene expression. Tachyzoite-specific genes were switched off, whereas bradyzoite-specific genes were switched on. The studies outlined above using microarrays [24–27] show that the controlling mechanisms behind stage conversion are more complex.

One of the most important outcomes of the study in Ref. [26] was the proposal of a model covering the changes in gene expression that occur during stage conversion. A cascade representing a hierarchy of gene expression was inferred from the detected patterns of expression during the study. The presentation of this model provides a working hypothesis upon which to base future studies on the molecular events that govern stage conversion in *T. gondii* and presumably other related cyst-forming coccidian.

Some of the key features of this model must also take into consideration other known facts of cell signaling that have been elucidated in *T. gondii*. It is also possible to speculate, based on our knowledge of the mechanisms of cell signaling in mammalian cells, what the controlling mechanisms behind cystogenesis are likely to be. Changes in gene expression are normally mediated by transcription factors in response to a complex series of cell-signaling events, which are initiated from interactions at the cell membrane. It is highly likely that the response of the tachyzoite to the initial stimulus to enter stage conversion is mediated by enzymes that produce second messengers [e.g. diacylglycerol (DAG) and cyclic nucleotides]. Available evidence shows that bradyzoite formation *in vitro* is associated with a transient increase in cAMP concentrations [37].

Given that the process of stage conversion is associated with a reduction in growth rate, it is also necessary to invoke hypotheses that the molecular events of stage conversion and the control of replication during the cell cycle are linked. Knowledge on the cell cycle and its control in *T. gondii* is sparse. The tachyzoite cell cycle is known to contain G1, S and M phases, with little or no G2 phase [38]. In addition, a cdc2 kinase has been reported in *T. gondii* [39] and cdc2 was subsequently recognized as a key central controlling element in the *T. gondii* cell cycle, as is the case for most eukaryotes [40]. The activity of cdc2 kinases is normally regulated by a cyclin subunit, of which several are recognized. To date, only one has been described in *T. gondii* [41].

An essential role for signaling through protein kinases must also be recognized. Protein phosphorylation and/or dephosphorylation normally play a central role in cell differentiation and proliferation, and in the regulation of metabolic pathways in all eukaryotic cells. Protein kinases and phosphatases have been widely studied in many cell types, but little knowledge is available on their presence and function in parasitic protozoa. Several protein kinases have now been described in *T. gondii* and, although the complete repertoire has yet to be described, available evidence suggests that they play a central role in stage conversion. For example, phorbol 12-myristate 13-acetate (PMA) enhances cyst formation from bradyzoites *in vitro* [42]. Although agonists and inhibitors of protein kinases are generally nondiscriminatory in their site of action [43,44], PMA is known biochemically to mimic the activity of natural second messengers, such as DAG [45,46], which exert their effect through modulating protein kinase activity.

A mitogen-activated protein (MAP) kinase activity was demonstrated in *T. gondii* whose activity is stimulated by PMA treatment of tachyzoites [47]. MAP kinases, as the name suggests, are normally regulated by membrane-based signal transduction events leading to a cascade of kinase activities, typically referred to as the MAP kinase pathway. Ultimately, signaling through the MAP kinase pathway leads to the activation or inactivation of transcription factors in the nucleus, thereby leading to changes in gene expression. A 14-3-3 homolog exists in the cyst-forming coccidia, and in other eukaryotes this is believed to be involved in upstream signaling events leading to MAP kinase activation [48]. In mammalian cells, shaggy/glycogen synthase kinase-3 kinase is located downstream of the MAP kinase signaling cascade [49] and a homolog of this kinase is known to exist in *T. gondii*.

Inhibition of cell growth by quercetin also correlates with induction of cyst formation [50]. Although this phenomenon is mediated by changes in expression of heat-shock protein [51], it has been recognized for some time in eukaryotes that the primary mode of action of quercetin is probably at a protein kinase activity.

Conclusions

The cyst-forming coccidia are an important group of pathogens and, consequently, the generation of new knowledge on the molecular mechanisms of stage conversion, and specifically cyst-formation, is important. Large-scale sequencing projects of the *T. gondii* transcriptome have enabled the generation and application of microarrays for the study of stage conversion. This in turn has opened 'new doors' for the investigation of the mechanisms that underpin the morphological and structural changes observed during the formation of tissue cysts. Experimental design and the application of appropriate data analysis methodologies need to be carefully considered if microarrays are to be correctly exploited. Nevertheless, the wide application of these tools in parasitology represents an important advance for this discipline.

Acknowledgements

We thank U. Singh and J. Boothroyd for providing microarray datasets.

References

1 Gross, U. *et al.* (1997) Reactivation of chronic toxoplasmosis: is there a link to strain-specific differences in the parasite? *Behring Inst. Mitt.*, 97–106
2 Soete, M. *et al.* (1994) Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of *Toxoplasma gondii in vitro*. *Exp. Parasitol.* 78, 361–370
3 Weiss, L.M. *et al.* (1995) A cell culture system for study of the development of *Toxoplasma gondii* bradyzoites. *J. Eukaryot. Microbiol.* 42, 150–157
4 Bohne, W. *et al.* (1994) Reduced replication of *Toxoplasma gondii* is necessary for induction of bradyzoite-specific antigens: a possible role for nitric oxide in triggering stage conversion. *Infect. Immun.* 62, 1761–1767
5 Weiss, L.M. *et al.* (1998) Bradyzoite development in *Toxoplasma gondii* and the hsp70 stress response. *Infect. Immun.* 66, 3295–3302
6 Lindsay, D.S. *et al.* (1998) Decoquinate induces tissue cyst formation by the RH strain of *Toxoplasma gondii*. *Vet. Parasitol.* 77, 75–81
7 Miller, C.M. *et al.* (1999) Cytokines, nitric oxide, heat shock proteins and virulence in *Toxoplasma*. *Parasitol. Today* 15, 418–422
8 Manger, I.D. *et al.* (1998) The surface of *Toxoplasma* tachyzoites is dominated by a family of glycosylphosphatidylinositol-anchored antigens related to SAG1. *Infect. Immun.* 66, 2237–2244
9 Lekutis, C. *et al.* (2001) Surface antigens of *Toxoplasma gondii*: variations on a theme. *Int. J. Parasitol.* 31, 1285–1292
10 Bohne, W. *et al.* (1999) Stage differentiation of the protozoan parasite *Toxoplasma gondii*. *Immunobiology* 201, 248–254
11 Weiss, L.M. and Kim, K. (2000) The development and biology of bradyzoites of *Toxoplasma gondii*. *Front. Biosci.* 5, D391–D405
12 Soete, M. and Dubremetz, J.F. (1996) *Toxoplasma gondii*: kinetics of stage-specific protein expression during tachyzoite–bradyzoite conversion *in vitro*. *Curr. Top. Microbiol. Immunol.* 219, 76–80
13 Bohne, W. *et al.* (1996) Bradyzoite-specific genes. *Curr. Top. Microbiol. Immunol.* 219, 81–91
14 Tomavo, S. (2001) The differential expression of multiple isoenzyme forms during stage conversion of *Toxoplasma gondii*: an adaptive developmental strategy. *Int. J. Parasitol.* 31, 1023–1031
15 Boothroyd, J.C. *et al.* (2003) DNA microarrays in parasitology: strengths and limitations. *Trends Parasitol.* 19, 470–476
16 Baldi, P. and Hatfield, G.W. (2002) *DNA Microarrays and Gene Expression: from Experiments to Data Analysis and Modeling*, Cambridge University Press
17 Knudsen, G.W. (2002) *A Biologist's Guide to Analysis of DNA Microarray Data*, John Wiley & Sons
18 Cummings, C.A. and Relman, D.A. (2000) Using DNA microarrays to study host–microbe interactions. *Emerg. Infect. Dis.* 6, 513–525
19 Ajioka, J.W. *et al.* (1998) Gene discovery by EST sequencing in *Toxoplasma gondii* reveals sequences restricted to the Apicomplexa. *Genome Res.* 8, 18–28
20 Manger, I.D. *et al.* (1998) Expressed sequence tag analysis of the bradyzoite stage of *Toxoplasma gondii*: identification of developmentally regulated genes. *Infect. Immun.* 66, 1632–1637
21 Ellis, J.T. *et al.* (2000) *Neospora caninum*: gene discovery through analysis of expressed sequence tags. In *A European Perspective on*

Neospora Caninum (Vol. 30) (Hemphill, A. and Gottstein, B., eds), pp. 909–913, *Int. J. Parasitol*.
22 Li, L. *et al.* (2003) Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. *Genome Res.* 13, 443–454
23 Buxton, D. and Innes, E.A. (1995) A commercial vaccine for ovine toxoplasmosis. *Parasitology* 110, S11–S16
24 Morrison, D.A. and Ellis, J.T. (2003) The design and analysis of microarray experiments: applications in parasitology. *DNA Cell Biol.* 6, 357–394
25 Cleary, M.D. *et al.* (2002) *Toxoplasma gondii* asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryot. Cell* 1, 329–340
26 Singh, U. *et al.* (2002) Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol. Microbiol.* 44, 721–733
27 Matrajt, M. *et al.* (2002) Identification and characterization of differentiation mutants in the protozoan parasite *Toxoplasma gondii*. *Mol. Microbiol.* 44, 735–747
28 Ouarzane, M. *et al.* (1998) *Eimeria tenella*: cloning and characterization of cDNA encoding a S3a ribosomal protein. *Gene* 225, 125–130
29 Oura, C.A. *et al.* (2002) Identification of a 40S Ribosomal protein (S17) that is differentially expressed between the macroschizont and piroplasm stages of *Theileria annulata*. *Int. J. Parasitol.* 32, 73–80
30 Mamoun, C.B. *et al.* (2001) Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite *Plasmodium falciparum* revealed by microarray analysis. *Mol. Microbiol.* 39, 26–36
31 Yang, S. and Parmley, S.F. (1997) *Toxoplasma gondii* expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts. *Gene* 184, 1–12
32 Knoll, L.J. and Boothroyd, J.C. (1998) Molecular biology's lessons about *Toxoplasma* development: stage-specific homologs. *Parasitol. Today* 14, 490–493
33 Gastens, M.H. and Fischer, H.G. (2002) *Toxoplasma gondii* eukaryotic translation initiation factor 4A associated with tachyzoite virulence is down-regulated in the bradyzoite stage. *Int. J. Parasitol.* 32, 1225–1234
34 Weiss, L.M. *et al.* (1999) The *in vitro* development of *Neospora caninum* bradyzoites. *Int. J. Parasitol.* 29, 1713–1723
35 Vonlaufen, N. *et al.* (2002) Exogenous nitric oxide triggers *Neospora caninum* tachyzoite-to-bradyzoite stage conversion in murine epidermal keratinocyte cell cultures. *Int. J. Parasitol.* 32, 1253–1265
36 Vonlaufen, N. *et al.* (2004) *In vitro* induction of *Neospora caninum* bradyzoites in vero cells reveals differential antigen expression, localization, and host-cell recognition of tachyzoites and bradyzoites. *Infect. Immun.* 72, 576–583
37 Kirkman, L.A. *et al.* (2001) Cyclic nucleotide signaling in *Toxoplasma gondii* bradyzoite differentiation. *Infect. Immun.* 69, 148–153
38 Radke, J.R. *et al.* (2001) Defining the cell cycle for the tachyzoite stage of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* 115, 165–175
39 Wastling, J.M. and Kinnaird, J.H. (1998) Isolation and characterization of a genomic clone encoding a cdc2-related kinase of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* 94, 143–148
40 Khan, F. *et al.* (2002) Cyclin-dependent kinase TPK2 is a critical cell cycle regulator in *Toxoplasma gondii*. *Mol. Microbiol.* 45, 321–332
41 Kvaal, C.A. *et al.* (2002) Isolation of a *Toxoplasma gondii* cyclin by yeast two-hybrid interactive screen. *Mol. Biochem. Parasitol.* 120, 187–194
42 Ricard, J. *et al.* (1996) TNF-alpha enhances *Toxoplasma gondii* cyst formation in human fibroblasts through the sphingomyelinase pathway. *Cell. Signal.* 8, 439–442
43 Casnellie, J.E. (1991) Protein kinase inhibitors: probes for the functions of protein phosphorylation. *Adv. Pharmacol.* 22, 167–205
44 Cuenda, A. and Alessi, D.R. (2000) Use of kinase inhibitors to dissect signaling pathways. *Methods Mol. Biol.* 99, 161–175
45 Castagna, M. *et al.* (1982) Direct activation of calcium-activated, phospholipid-dependent protein kinase by tumor-promoting phorbol esters. *J. Biol. Chem.* 257, 7847–7851
46 Kikkawa, U. *et al.* (1984) Protein phosphorylation and mechanism of action of tumor-promoting phorbol esters. *Adv. Cyclic Nucleotide Protein Phosphorylation Res.* 17, 437–442
47 Roisin, M.P. *et al.* (2000) Biochemical characterization of mitogen-activated protein (MAP) kinase activity in *Toxoplasma gondii*. *Parasitol. Res.* 86, 588–598
48 Lally, N.C. *et al.* (1996) Evaluation of two *Neospora caninum* recombinant antigens for use in an enzyme-linked immunosorbent assay for the diagnosis of bovine neosporosis. *Clin. Diagn. Lab. Immunol.* 3, 275–279
49 Qin, C.L. *et al.* (1998) Cloning and *in vitro* expression of TPK3, a *Toxoplasma gondii* homologue of shaggy/glycogen synthase kinase-3 kinases. *Mol. Biochem. Parasitol.* 93, 273–283
50 Halonen, S.K. *et al.* (1999) Effect of cytokines and quercetin on *Toxoplasma gondii* cyst induction in murine astrocytes. *J. Eukaryot. Microbiol.* 46, 83S–84S
51 Dobbin, C.A. *et al.* (2002) Heat shock protein 70 is a potential virulence factor in murine *Toxoplasma* infection via immunomodulation of host nuclear factor kappa beta and nitric oxide. *J. Immunol.* 169, 958–965
52 Schena, M. (1996) Genome analysis with gene expression microarrays. *Bioessays* 18, 427–431
53 Stears, R.L. *et al.* (2003) Trends in microarray analysis. *Nat. Med.* 9, 140–145
54 Gail, M. *et al.* (2001) Transcriptional profile of *Toxoplasma gondii*-infected human fibroblasts as revealed by gene-array hybridization. *Mol. Genet. Genomics.* 265, 905–912
55 Blader, I.J. *et al.* (2001) Microarray analysis reveals previously unknown changes in *Toxoplasma gondii*-infected human cells. *J. Biol. Chem.* 276, 24223–24231
56 Manger, I.D. and Relman, D.A. (2000) How the host 'sees' pathogens: global gene expression responses to infection. *Curr. Opin. Immunol.* 12, 215–218
57 Staudt, L.M. and Brown, P.O. (2000) Genomic views of the immune system. *Annu. Rev. Immunol.* 18, 829–859
58 Huang, Q. *et al.* (2001) The plasticity of dendritic cell responses to pathogens and their components. *Science* 294, 870–875
59 Boldrick, J.C. *et al.* (2002) Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 99, 972–977
60 Hoffmann, K.F. *et al.* (2002) Identification of *Schistosoma mansoni* gender-associated gene transcripts by cDNA microarray profiling. *Genome Biol.* 3, RESEARCH0041.
61 Diehl, S. *et al.* (2002) Analysis of stage-specific gene expression in the bloodstream and the procyclic form of *Trypanosoma brucei* using a genomic DNA-microarray. *Mol. Biochem. Parasitol.* 123, 115–123
62 Hayward, R.E. *et al.* (2000) Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol. Microbiol.* 35, 6–14
63 Ganesan, K. *et al.* (2002) Stochastic versus stable transcriptional differences on *Plasmodium falciparum* DNA microarrays. *Int. J. Parasitol.* 32, 1543–1550