

An evaluation of serial analysis of gene expression (SAGE) in the parasitic nematode, *Haemonchus contortus*

P. J. SKUCE*, R. YAGA, F. A. LAINSON and D. P. KNOX

Moredun Research Institute, Pentlands Science Park, Bush Loan, Penicuik, Midlothian EH26 0PZ, Scotland, UK

(Received 5 October 2004; revised 3 November 2004; accepted 3 November 2004)

SUMMARY

This study evaluated a relatively new molecular technique, serial analysis of gene expression (SAGE), as a tool for quantifying gene expression in the ovine abomasal nematode *Haemonchus contortus* for which there is relatively limited (~20% gene coverage) sequence information. SAGE technology generates data that are both qualitative and quantitative and, as such, compliments other functional genomics approaches such as EST analysis and micro-array. Prior to embarking on large-scale comparisons, the present study was initiated to establish (i) how well SAGE and EST data taken from the same life-cycle stage would compare, (ii) how easily SAGE tags could be assigned to genes given that the genome sequence is not available and (iii) whether it would be possible to extend the sequences of the SAGE tags to facilitate their identification. Of 2825 tag sequences analysed from adults harvested 28 days post-infection, the identity of the encoding gene could be ascribed to 63% of the tags. The relative abundance of these genes, arbitrarily categorized on the basis of function, was comparable with that of an EST dataset also from adults ($n=2317$). In addition, tag sequences could be readily extended and thereby identified using a tag-based primer and Reverse Transcription-PCR.

Key words: *Haemonchus contortus*, serial analysis of gene expression, SAGE.

INTRODUCTION

Now that we are entering the so-called “post-genome era” (Nowak, 1995), it is anticipated that detailed analyses of gene expression in parasites may help identify gene products that are critical for parasite survival and that may prove to be novel drug or vaccine targets (Knox, 2004). This situation has become particularly acute for the gastrointestinal nematodes due to the inexorable rise in drug resistant parasites and the fact that there are no anti-nematode vaccines near market (e.g. Knox *et al.* 2003). Such analyses would involve, for example, identifying genes that are expressed in either a parasite-, stage- or tissue-specific manner or those which are up-regulated in response to drug treatment or vaccination. To date, comparative studies have employed methodologies such as differential colony/plaque hybridization, suppressive-subtractive hybridization, mRNA differential display and Real-time PCR (Knox, 2004). These methods have their respective advantages and disadvantages but a limitation is that they can examine a small number of genes at a time. DNA micro-array technology can extend such studies to a genome-wide scale but this potential is restricted by the lack of defined sequence data available for most of the important gastrointestinal species. Considerable research effort has been invested into the nematode expressed sequence tag

(EST) project (Parkinson *et al.* 2003) with approximately 400 000 nematode ESTs on the web at NEMBASE (<http://www.nematodes.org>) and at www.ncbi.nlm.nih.gov/dbEST/, including data-sets from the 3 major gastrointestinal species, namely *Haemonchus contortus*, *Ostertagia ostertagi* and *Teladorsagia circumcincta*. The largest data-set is for *Haemonchus*; at present there are some 17 269 available ESTs which cluster into a total of 4145 contiguous sequences. However, these represent ~20% coverage of the likely ~20 000 genes encoded within the genome of a nematode (*C. elegans* Sequencing Consortium, 1998), a fact which would potentially compromise the output from array analysis of gene expression. However, both the nematode EST initiative and a small-scale EST survey of L1, L3, L5 and adult *Haemonchus* (Hoekstra *et al.* 2000) have shown a distinct shift in gene expression from house-keeping genes to less well-conserved genes encoding putative extracellular proteins. More recently, an RNA arbitrarily-primed PCR (RAP-PCR) approach was applied to the study of developmentally regulated genes in *H. contortus* (Hartman *et al.* 2001) and revealed that approximately one third of the genes expressed at any given stage were unique to that stage. The much larger data-sets currently maintained at NEMBASE have been set up such that they can be interrogated *in silico* on the basis of sequence annotation, similarity and stage-specificity of expression (Parkinson *et al.* 2003). Expression can be quantified on the basis of EST frequency in the data-sets. For example, >10% of all the ESTs for

* Corresponding author. Tel: +44 (0) 131 445 5111. Fax: +44 (0) 131 445 6111. E-mail: skucp@mri.sari.ac.uk

Haemonchus are from a small family of genes that shares only limited similarity with hypothetical genes in *C. elegans* (termed F54D5). This gene appears to be expressed exclusively in the adult and there are no close homologues in *Ostertagia* or *Teladorsagia*.

It would be prohibitively expensive to undertake EST sequencing projects to detect genes that are transiently expressed by a particular parasite stage in response to, for example, developing host immunity, experiments which could be conducted if the entire genome was available for array analyses. An alternative method, called serial analysis of gene expression (SAGE, Velculescu *et al.* 1995) offers the potential to perform such comparisons in a much more cost-effective manner. The method employs standard molecular biology techniques and should be within the capability of an experienced molecular parasitology laboratory. Moreover, the method can be adapted for use with limited amounts of starting material (Ye *et al.* 2000), often a problem when working with helminth parasites. The SAGE method is based on 3 key principles (Velculescu *et al.* 1995). Firstly, a short (9–14 bp) sequence tag contains sufficient information to identify the transcript from which it originates. Secondly, when cloned and concatenated, these tags can be sequenced in a much more efficient manner and, thirdly, the number of times a given tag is observed equates to the expression level of its original transcript. SAGE has been widely used in studies on human cancer, immunology, physiology and developmental biology (Tuteja & Tuteja, 2004). SAGE has recently been applied to a number of different organisms including the protozoan parasites *Plasmodium falciparum* (Munasinghe *et al.* 2001) and the free-living nematode, *Caenorhabditis elegans* (Jones *et al.* 2001). A fully annotated genome sequence exists for all of these species and this has been assumed to be a prerequisite for SAGE analysis. In the present study, we have attempted to evaluate SAGE as a tool for gene expression profiling in *Haemonchus*. Prior to embarking on large-scale comparisons, we initially needed to establish (i) how well SAGE and EST data taken from the same life-cycle stage would compare, (ii) how easily SAGE tags could be assigned to genes given that the genome sequence is not available and (iii) whether it would be possible to extend the sequences of unknown SAGE tags in order to facilitate their identification.

MATERIALS AND METHODS

Construction of SAGE library

A SAGE library was constructed from adult (28-day-old) *Haemonchus* using a commercial I-SAGE kit (Invitrogen), essentially according to the manufacturer's instructions. Briefly, double-stranded cDNA was synthesized from ~100 ng mRNA using

a biotinylated oligo-dT primer. The cDNA was subdivided into 2 pools (A and B) and the 3' ends of the respective cDNAs were isolated by binding to streptavidin-coated magnetic beads and cleaved with the 'anchoring enzyme' *Nla*III. The exposed 5' ends were then blunted with Klenow DNA polymerase and ligated on to their respective double-stranded adaptors (A or B). The adaptors contain primer-binding sites for subsequent PCR amplification and a type II restriction enzyme cleavage site (in this case *Bsm*FI) for release of the individual tag sequences. Following *Bsm*FI cleavage, tags were concatamerized using T4 DNA ligase and cloned into the pZero plasmid (Invitrogen). An aliquot of the finished library was plated onto LBzeocin plates containing X-gal and IPTG. A number of the resultant white colonies were picked at random into 100 μ l of H₂O in a microtitre plate, 5 μ l of this suspension were used as template in an M-13 Forward/Reverse primer PCR. Reaction conditions were set up with final primer and dNTP concentrations of 0.1 μ M and 0.2 mM, respectively such that PCR products could be sequenced directly. Samples were denatured initially for 3 min at 95 °C followed by 39 cycles of 94 °C for 1 min; 54 °C for 1 min; 72 °C for 2 min with a final cycle of 94 °C for 30 s; 54 °C for 10 min; 72 °C for 10 min and cooled to 4 °C. PCR products were examined on a 0.8% agarose gel prior to being sequenced with the M-13 Forward primer on an ABI Prism 377 automated sequencer.

Sequence analyses

The primary objective was to sequence a similar number of tags to the number of ESTs derived from the same life-cycle stage maintained at NEMBASE (28 days old adult, Library # 8631; 2317ESTs). Tag sequences ($n=2825$) were subjected to analysis using the e-SAGE programme (available online at <http://iubio.bio.indiana.edu/soft/molbio/nhgri/eSAGE>), which removes vector sequences and sorts and counts valid tags. The e-SAGE software was run on a UNIX platform and customised such that individual tags could be BLAST searched locally against the complete *Haemonchus* data-set in the GenBank database.

Tag-based Reverse Transcriptase-PCR

Briefly, an antisense oligonucleotide complementary to the SAGE tag itself was synthesized with multiple (>5) inosines at the 5' end to enhance stability and primer binding, and employed in a PCR reaction in combination with a suitable vector primer (in this case T7) and an aliquot of an 11-day-old *Haemonchus* cDNA library in lambda ZAP-XR as a template. PCR conditions used were 94 °C for 5 min, 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min for 30 cycles

Table 1. Primer sequences used in the study

(MEP3, PEP1, Thrombo and HMCP4 refer to specific target genes from *Haemonchus contortus* available from GenBank with Accession numbers AF08172, Z72490, AF043121 and Z69345, references Smith *et al.* (2000), Longbottom *et al.* (1997), Skuce *et al.* (2001) and Skuce *et al.* (1999) respectively.)

Gene	Gene-specific antisense primer
MEP3SAGE	5' IIIIATTTACTCGGCATG 3'
PEP1SAGE	5' IIIITCGTATTCCTCATG 3'
ThromboSAGE	5' IIIIIGAACCTTCTCACATG 3'
HMCP4SAGE	5' IIIIATCTTAATAGCATG 3'
T7 primer	5' TAATACGACTCACTATAGGGC 3'

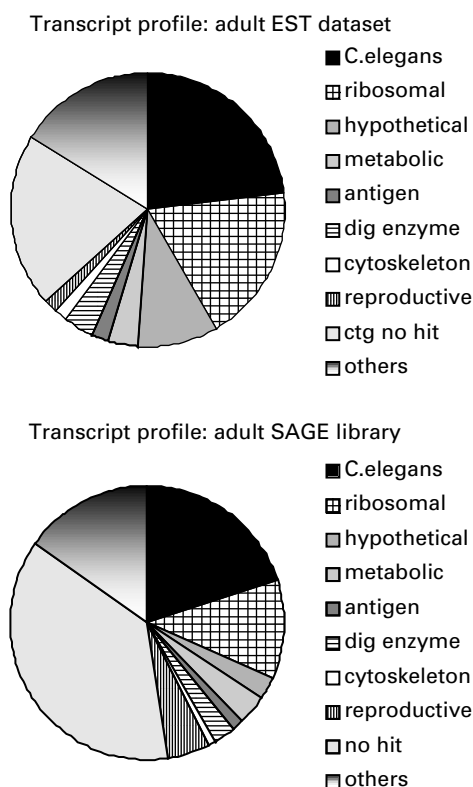


Fig. 1. A comparison of the relative abundance (%) of ESTs (library 8613) and SAGE tags grouped on the basis of function.

followed by a final extension cycle at 72 °C for 7 min. The primer sequences are given in Table 1.

RESULTS

A total of 2825 SAGE tags representing 1214 individual species was identified in the present study. Only tags producing a 100% match over the full 14 bp were scored as hits and only tags with an abundance score of >1 were included in the analysis (to discount possible sequencing errors). The identity of the most abundant tags was compared with the equivalent from adult EST analyses of cDNA libraries prepared from USA- and UK-derived worm populations (Table 2), the latter being harvested

from the host at the same time-point as those used for SAGE analysis. The table shows the top 25 tags ranked on the basis of abundance compared EST data for similar populations of UK-derived (Library 8613) and USA-derived worms (library 12015; both available at NEMBASE, www.nematodes.org). The striking aspect of this comparison is that at first glance there appeared to be very little relationship between gene abundances in the datasets. However, on closer inspection, some similarities were apparent, particularly between the SAGE data and the EST data from UK parasites (8613). Amongst the most abundant genes represented in both datasets (results expressed as % abundance in SAGE: % abundance in EST) are the hypothetical *C. elegans* proteins F54D5.3 (2.4:7.9), F58G1.4 (1.1:0.9), ribosomal proteins (4.2:5.0), heat shock protein homologues (0.7:0.5) and cysteine proteinases (0.4:0.5).

A more detailed analysis of the transcript profiles generated from the EST and SAGE data, in which the respective genes are classified on the basis of likely function, revealed essentially the same pattern of gene expression between the two methods, the only significant discrepancy being the proportion of SAGE tags for which there is no hit in the database search (Fig. 1). A total of 37% of the SAGE tags compared to 20.7% ESTs were in this category in the present analysis (Fig. 1).

A comparison of the relative abundance of some genes of particular interest in the authors' laboratory in the EST (Library 8613) and SAGE datasets was undertaken (Table 3, Knox *et al.* 2003). In agreement with the general analysis described, the relative abundance of the genes identified was similar in the two datasets. For example, cathepsin B and metalloproteinases were particularly abundant whereas thrombospondin and cathepsin L were found at low frequency in both data-sets.

It is important to be able to identify unknown tags and here a tag-based RT-PCR (van den Berg *et al.* 1999) was evaluated for this purpose. This was tested using primers designed to the SAGE tags predicted for 4 genes of interest, namely MEP3 and HMCP4 (Fig. 2), PEP1 and thrombospondin. MEP3 and

Table 2. The 25 most abundant SAGE tags compared with similar data extracted from EST data-sets for USA- and UK-derived *Haemonchus contortus* populations (EST libraries available at NEMBASE (www.nematodes.org))

SAGE Library (No. of tags = 2438)			USA adult library 12015 (n = 875)			UK adult EST library 8613 (n = 2317)		
No. of tags	% Total	Identity	No. of ESTs	% Total	Identity	No. of ESTs	% Total	Identity
102	4.2	18S ribosomal protein	91	10.4	<i>H. contortus</i> cysteine proteinase	162	7.0	<i>C. elegans</i> protein, F54D5.3
60	2.5	Vitellogenin	28	3.2	<i>C. elegans</i> protein, F35C5.6	84	3.6	60S Ribosomal protein
46	1.9	<i>C. elegans</i> protein, F54D5.3	24	2.7	No hit	32	1.4	RIKEN (RNA binding protein)
41	1.7	Hyohtetical <i>C. elegans</i> protein	21	2.4	No hit	23	1.0	C-type lectin precursor
31	1.3	Hypothetical protein	19	2.2	Putative GTP-binding protein	20	0.9	<i>C. elegans</i> protein, F54D5.3
28	1.1	<i>C. elegans</i> F58G1.4 protein	14	1.6	No hit	20	0.9	<i>C. elegans</i> F58G1.4 protein
26	1.1	ATP synthetase	13	1.5	<i>C. elegans</i> protein, F41C3.4	16	0.7	No hit
23	1.0	Hypoxanthine phosphoribosyl transferase	11	1.3	No hit	16	0.7	No hit
21	0.9	Cytochrome C oxidase	11	1.3	Adenosyl methioine decarboxylase	14	0.6	HIV envelope glycoprotein
19	0.8	No hit	11	1.3	No hit	14	0.6	<i>C. elegans</i> protein, F54D5.3
18	0.8	Lipid transfer protein	10	1.1	<i>H. contortus</i> Cysteine proteinase	14	0.6	<i>Haemonchus</i> glutamate dehydrogenase
15	0.6	Purine nucleoside phosphorylase	10	1.1	HIV envelope glycoprotein	11	0.5	<i>C. elegans</i> protein, F54D5.3
12	0.5	Elongation factor	9	1.0	Hypothetical protein	11	0.5	Cysteine proteinase
12	0.5	No hit	8	0.9	No hit	11	0.5	Heat shock protein
12	0.5	NADH dehydrogenase subunit	8	0.9	<i>C. elegans</i> Hypothetical protein	11	0.5	No hit
11	0.5	<i>C. elegans</i> F54 D5.3/4	8	0.9	<i>C. elegans</i> Hypothetical protein	10	0.4	Hypothetical protein
10	0.4	<i>H. contortus</i> cysteine proteinase	8	0.9	<i>Ascaris thymidylate synthetase</i>	10	0.4	Heat shock protein
10	0.4	<i>H. contortus</i> 26S rRNA gene	8	0.9	No hit	10	0.4	No Hit
9	0.4	<i>Haemonchus</i> ESTs	7	0.8	<i>Haemonchus cathepsin L</i>	10	0.4	No hit
9	0.4	<i>Haemonchus</i> ESTs	7	0.8	No hit	10	0.4	Major sperm protein
9	0.4	<i>Haemonchus</i> ESTs	7	0.8	<i>C. elegans</i> Hypothetical protein	9	0.4	Hypothetical protein
8	0.3	<i>Haemonchus</i> ESTs	7	0.8	<i>C. elegans</i> Hypothetical protein	9	0.4	No hit
8	0.3	<i>Haemonchus</i> ESTs	7	0.8	<i>C. elegans</i> Hypothetical protein	9	0.4	<i>C. elegans</i> JC8.8 protein
8	0.3	<i>Haemonchus</i> ESTs	7	0.8	Acetyl CoA transferase	9	0.4	No hit
8	0.3	Heat shock protein	6	0.7	No hit	9	0.4	No hit

Table 3. The relative abundance of specific genes of interest in the EST and SAGE data-sets

	% Abundance	
	ESTs	SAGE
Thrombospondin	0.027	0.041
Pepsinogen	0.362	0.013
MEPs	1.02	0.656
Galectin	0.026	0.041
Cystatin	0.1	0.082
Cat B	5.813	1.189
Cat L	0.013	0.012
GDH	0.037	0.016
SOD	0.168	0.082

HMCP4 are members of closely-related, multi-gene families in *Haemonchus*, whereas PEP1 and thrombospondin are encoded by single-copy genes (Smith *et al.* 2000; Skuce *et al.* 1999, 2001). Without exception, the sequences obtained from the RT-PCR products were as predicted for the respective genes, thereby demonstrating the utility of this approach.

DISCUSSION

The present study evaluated SAGE as a tool for gene expression profiling in *Haemonchus*. This was addressed by comparing gene abundances in SAGE and EST data-sets taken from the same life-cycle stage, evaluating how easily SAGE tags could be assigned to genes given that the *Haemonchus* genome sequence is not available and testing a tag-based RT-PCR to extend the sequences of unknown SAGE tags in order to facilitate their identification.

While there was some agreement between the SAGE tag dataset and the UK-derived EST data, the apparent lack of similarity between the USA and UK parasite-derived data-sets was surprising (Table 2). For example, F54D5 proteins are much less abundant in the USA dataset and were not detected in the previous, limited EST-based survey of gene expression using a Netherlands strain of this parasite (Hoekstra *et al.* 2000). This may simply represent between-library variation, subtle temporal changes in gene expression or more fundamental geographical strain differences. The former is unlikely because the F54D5 proteins are also highly represented in another UK library (Library 8396, NEMBASE, www.nematodes.org) sampled in the EST survey, although this library was prepared from worms harvested at 11 days post-infection. The latter possibilities could be readily addressed by conducting SAGE comparisons of gene abundance in more than one library prepared from the same population of worms. Another point worthy of consideration

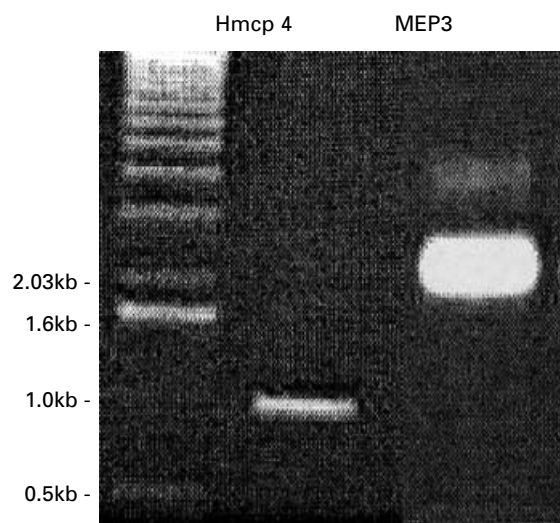


Fig. 2. PCR verification of SAGE specificity. SAGE tags were predicted from 4 known genes, two from multigene families, and these tag sequences were employed as antisense primers in an anchored PCR using a *Haemonchus contortus* cDNA library as template. PCR products were cloned and sequenced and without exception proved to be the targeted gene. The figure shows the amplification products obtained for hmcp 4 and MEP3.

is that most of the genes listed individually are represented at a low level ($\sim 2\%$).

The proportion of SAGE tags (37%) for which no homologues were identified was higher than the corresponding figure for the UK-derived ESTs (21%). This discrepancy could have a number of explanations. Either the SAGE tag represents a genuinely novel gene or its corresponding EST sequence does not contain the 3' terminus where the tag is located. EST sequences are typically in the order of ~ 500 bp in length so the probability of generating a significant match in a homology search is relatively high. Nonetheless, a significant proportion of nematode ESTs have no match in the databases, equating to 20% for the adult library profiled in the present study. Since SAGE tags are short and typically originate from the 3' end of most transcripts (whereas most ESTs are sequenced from the 5' end) a further proportion may not match sequences in a database search. Hence, the utility of the tag-based PCR for extension of specific sequences of interest was evaluated here and discussed in the following.

Antisense oligonucleotides complementary to the SAGE tags for several genes of interest in the authors' laboratory were synthesized with multiple (>5) inosines at the 5' end to enhance stability and primer binding. The primers were then used in combination with a vector primer to extend the tag sequence in the 5' direction by PCR, using a cDNA library lysate as template. All PCRs resulted in products encoding the specific target gene despite the

fact that some (MEP3 and HMCP4) were directed at genes that were members of gene families.

The number of SAGE tags sequenced is an important consideration and has important implications for the cost-effectiveness and scale of comparative analyses. One might assume that the more tags that are sequenced, the more representative the gene expression profile. However, this does not necessarily appear to be the case. For example, in a recent study of gene expression in HeLa cells, a total of 80 000 tags were sequenced and the transcript profile compared with profiles generated from smaller subsets representing 2000, 4000, 10 000, 20 000 and 40 000 tags (Yamamoto *et al.* 2001). Comparisons between subsets revealed practically no difference in tag proportions for highly abundant genes, with discrepancies only being evident for tags with an abundance of <3 in the smaller subsets (Yamamoto *et al.* 2001). Similarly, in the first study of its kind, a SAGE library comprising 4606 tags was used to reveal drug-induced changes in gene expression in the asexual blood forms of *P. falciparum* (Munasinghe *et al.* 2001). Although approximately 2500 tags were sequenced in the present analysis, the gene expression profile generated was comparable with that of the adult EST data-set and could be generated at a fraction of the cost and effort. Such findings indicate that SAGE, even on a relatively small scale, can be applied to analyses of gene expression in parasitic nematodes where database sequence information is limited. Moreover, SAGE could be used to analyse gene expression in parasites exposed to an environmental stress such as drug therapy or changes in gene expression induced by the host immune response. It could be anticipated that elevated expression of novel parasite genes, which may facilitate survival, would be detected by the unique appearance of novel tags when compared to tag populations from parasites harvested from parasite-naïve hosts. In the case of the EST datasets available, for example, for *Haemonchus*, all parasites used were harvested from immunologically naïve donor lambs and, hence should not contain ESTs for genes which are transiently expressed as a response to host immune effectors. The ability to extend the tag sequences in a gene-specific manner using tag-based RT-PCR, as demonstrated here, will enhance the applicability of tag-derived homology searches.

Obviously, the full potential of SAGE should be realised when a completely annotated genome is made available for *Haemonchus*. Even now it represents an affordable, in-house alternative to EST and micro-array analyses.

This work was generously funded by the Scottish Executive Environment and Rural Affairs Department. The authors are very grateful to Professor Peter O'Shaughnessy for helpful discussions before this work commenced.

REFERENCES

- THE *C. ELEGANS* SEQUENCING CONSORTIUM. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018.
- HOEKSTRA, R., VISSER, A., OTSEN, M., TIBBEN, J., LENSTRA, J. A. & ROOS, M. H. (2000). EST sequencing of the parasitic nematode *Haemonchus contortus* suggests a shift in gene expression during transition to the parasitic stages. *Molecular and Biochemical Parasitology* **110**, 53–68.
- HARTMAN, D., DONALD, D. R., NIKOLAOU, S., SAVIN, K. W., HASSE, D., PRESIDENTE, P. J. & NEWTON, S. E. (2001). Analysis of developmentally regulated genes of the parasite *Haemonchus contortus*. *International Journal for Parasitology* **31**, 1236–1245.
- LONGBOTTOM, D., REDMOND, D. L., RUSSELL, M., LIDDELL, S., SMITH, W. D. & KNOX, D. P. (1997). Molecular cloning and characterisation of a putative aspartate proteinase associated with a gut membrane protein complex from adult *Haemonchus contortus*. *Molecular and Biochemical Parasitology* **88**, 63–72.
- JONES, S. J., RIDDLE, D. L., POUZYREV, A. T., VELCULESCU, V. E., HILLIER, L., EDDY, S. R., STRICKLIN, S. L., BAILLIE, D. L., WATERSTON, R. & MARRA, M. A. (2001). Changes in gene expression associated with developmental arrest and longevity in *Caenorhabditis elegans*. *Genome Research* **11**, 1323–1324.
- KNOX, D. P. (2004). Technological advances and genomics in metazoan parasites. *International Journal for Parasitology* **34**, 139–152.
- KNOX, D. P., REDMOND, D. L., NEWLANDS, G. F., SKUCE, P. J., PETTIT, D. & SMITH, W. D. (2003). The nature and prospects for gut membrane proteins as vaccine candidates for *Haemonchus contortus* and other ruminant trichostrongyloids. *International Journal for Parasitology* **33**, 1129–1137.
- MUNASINGHE, A., PATANKAR, S., COOK, B. P., MADDEN, S. L., MARTIN, R. K., KYLE, D. E., SHOAIPI, A., CUMMINGS, L. M. & WIRTH, D. F. (2001). Serial analysis of gene expression (SAGE) in *Plasmodium falciparum*: application of the technique to A-T rich genomes. *Molecular and Biochemical Parasitology* **113**, 23–34.
- NOWAK, R. (1995). Entering the post genome era. *Science* **270**, 368–371.
- PARKINSON, J., MITREVA, M., HALL, N., BLAXTER, M. & McCARTER, J. P. (2003). 400 000 nematode ESTs on the Net. *Trends in Parasitology* **19**, 283–286.
- SKUCE, P. J., REDMOND, D. L., LIDDELL, S., STEWART, E. M., SMITH, W. D. & KNOX, D. P. (1999). Molecular cloning and characterisation of gut-derived cysteine proteinases associated with a host protective extract from *Haemonchus contortus*. *Parasitology* **119**, 396–405.
- SKUCE, P. J., NEWLANDS, G. F. J., STEWART, E. M., PETTIT, D., SMITH, S. K., SMITH, W. D. & KNOX, D. P. (2001). Cloning and characterisation of thrombospondin, a novel multidomain glycoprotein found in association with a host protective gut extract from *Haemonchus contortus*. *Molecular and Biochemical Parasitology* **117**, 241–244.
- SMITH, W. D., SMITH, S. K., PETTIT, D., NEWLANDS, G. F. & SKUCE, P. J. (2000). Relative protective properties of three membrane glycoprotein fractions from *Haemonchus contortus*. *Parasite Immunology* **22**, 63–71.
- TUTEJA, R. & TUTEJA, N. (2004). Serial analysis of gene expression: applications in human studies.

- Journal of Biomedicine and Biotechnology* **2004**, 113–120.
- VAN DEN BERG, A., VAN DER LEIJ, J. & POPPEMA, S. (1999). Serial analysis of gene expression: rapid RT-PCR analysis of unknown SAGE tags. *Nucleic Acids Research* **27**, 17.
- VELCULESCU, V. E., ZHANG, L., VOGELSTEIN, B. & KINZLER, K. W. (1995). Serial analysis of gene expression. *Science* **270**, 484–487.
- YAMAMOTO, M., WAKATSUKI, T., HADA, A. & RYO, A. (2001). Use of serial analysis of gene expression (SAGE) technology. *Journal of Immunological Methods* **250**, 45–66.
- YE, S. Q., ZHANG, L. Q., ZHENG, F., VIRGIL, D. & KWITEROVICH, P. O. (2000). MiniSAGE: Gene expression profiling using serial analysis of gene expression from 1 μ g total RNA. *Analytical Biochemistry* **287**, 144–152.