Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular & Biochemical Parasitology 141 (2005) 1–13
MOLECULAR & BIOCHEMICAL PARASITOLOGY

An oligonucleotide microarray for transcriptome analysis of *Schistosoma mansoni* and its application/use to investigate gender-associated gene expression☆

Jennifer M. Fitzpatrick[a], David A. Johnston[b], Gary W. Williams[c], Debbie J. Williams[c], Tom C. Freeman[c], David W. Dunne[a], Karl F. Hoffmann[a,*]

[a] *Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK*
[b] *Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*
[c] *Rosalind Franklin Centre for Genomics Research, Hinxton, Cambridge CB10 1SB, UK*

Received 14 September 2004; received in revised form 12 December 2004; accepted 4 January 2005
Available online 2 February 2005

Abstract

Global profiling transcriptomes of parasitic helminths offers the potential to simultaneously identify co-ordinately expressed genes, novel genetic programs and uniquely utilized metabolic pathways, which together provide an extensive and new resource for vaccine and drug discovery. We have exploited this post-genomic approach to fabricate the first oligonucleotide DNA microarray for gene expression analysis of the parasitic trematode *Schistosoma mansoni*. A total of 17,329 *S. mansoni* DNA sequences were used to design a microarray consisting of 7335 parasite elements or approximately 50% of this parasite's transcriptome. Here, we describe the design of this new microarray resource and its evaluation by extending studies into gender-associated gene expression in adult schistosomes. We demonstrate a high degree of reproducibility in detecting transcriptional differences among biologically replicated experiments and the ability of the microarray to distinguish between the expression of closely related gene family members. Importantly, for issues related to sexual dimorphism, labour division, gamete production and drug target discovery, 197 transcripts demonstrated a gender-biased pattern of gene expression in the adult schistosome, greatly extending the number of sex-associated genes. These data demonstrate the power of this new resource to facilitate a greater understanding into the biological complexities of schistosome development and maturation useful for identifying novel intervention strategies.

Keywords: *Schistosoma*; Helminth; DNA microarray; Gene expression

1. Introduction

Parasite expressed sequence tag (EST) and genomic sequencing projects have proven to be an invaluable resource for parasite gene discovery and have led to the identification of numerous putative gene products [1–3]. Despite the large availability of DNA sequence information, true potential of this resource will only be realized upon the assignment of gene function within an actual biological and cellular context, thus leading to the possible functional annotation of many important parasitic genomes. Towards this goal, several investigators have developed DNA microarrays to probe and begin to elucidate the role of specific gene products in the lifestyle, pathogenicity and fundamental biology of multiple parasites [4–6]. This approach, in combination with contin-

☆ Microarray data reported in this paper is available in the ArrayExpress database at EBI under the reference numbers A-MEXP-134 (description of *S. mansoni* oligonucleotide microarray) and E-MEXP-223 (all microarray data).

* Corresponding author. Tel.: +44 1223 333 338; fax: +44 1223 333 741.
E-mail address: kfh24@cam.ac.uk (K.F. Hoffmann).

0166-6851/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2005.01.007

ued wide ranging genomic and EST sequencing has brought together genomic and functional-genomic data to reveal new insights into the complex pathogenic parasitic lifestyle.

Gene discovery and functional analysis of schistosomes has also benefited from such an approach [5,7]. These studies have revealed cDNA microarrays to be particularly useful and reliable in identifying gender-associated transcripts in both *Schistosoma mansoni* [5] and *S. japonicum* [7]. However, while sampling small gene subsets proved fruitful for the initiation of these investigations, utilizing a high-throughput approach on a whole genome scale will dramatically increase the transcriptional understanding of parasite sexual biology where current thinking proposes that males and females evolved to maximize independent functional roles important to the survival of the parasite. Presently, the *S. mansoni* genome contains approximately 14,000 predicted genes, and like many other parasites, most of these genes display no database homology and therefore, have no functional annotation [8–10]. Assigning some putative function or association, based on expression profiling by DNA microarray analysis, may provide some of the most promising research areas for elucidating the molecular basis of parasite biology.

Here we describe the design, fabrication and validation of a new DNA microarray for schistosome transcriptome analysis based on the use of long oligonucleotide probes. Long oligonucleotide DNA microarrays were constructed as highly sensitive alternatives to cDNA microarrays due to inherent advantages throughout production and experimental use [11]. The DNA microarray contains 7335 *S. mansoni* probes covering approximately 50% of the total estimated gene complement and was used here to expand upon our previous studies of schistosome conjugal biology. The experiments described provide a dual function: (1) to specifically characterize the fabricated oligonucleotide DNA microarray allowing sensitive, reproducible gene expression results to be generated by multiple users and laboratories; and (2) to further elucidate the expression profiles of male and female adult parasites with the goal of expanding our knowledge relating to sexual maturation, sexual dimorphism, labour division and gamete production. Ultimately, investigations into *S. mansoni* transcriptional mechanisms will likely generate new insights into the development and maintenance of this helminth's dioescious lifestyle, leading to the identification of novel drug targets or vaccine candidates.

2. Materials and methods

2.1. Parasites

Adult male and female *S. mansoni* (NMRI Puerto Rican strain) were perfused from percutaneously infected mice at 7 weeks after challenge with independent batches of 250 cercariae each shed from albino *Biomphalaria glabrata*. After perfusion, both immature and mature worms were counted and sex-separated. Miracidia used to infect *B. glabrata* were hatched from eggs collected from mouse livers 7 weeks after infection [12].

2.2. Design of S. mansoni DNA oligonucleotide probes

S. mansoni DNA elements chosen for oligonucleotide design were selected from EST sequences available from public databases as of June 28, 2002, full-length mRNA and genomic DNA (gDNA) sequences (using NCBI Entrez limits excluding ESTs, STSs, GSSs, TPAs, patents and working drafts) available from public databases as of April 2003 and from one full-length mRNA sequence identified in our laboratory (AY267032—*S. mansoni* arginase). The 16,815 EST sequences were clustered using the CAP3 DNA sequence assembly program [13] into 2076 contigs (representing more than one EST sequence) and 5049 singletons (representing only one EST sequence) for a total of 7125 unique DNA sequence clusters. In addition to these 7125 non-redundant EST clusters, 513 full-length mRNA (some redundancy with respect to EST clusters) and gDNA sequences were included to bring the total number of DNA sequences used as templates for oligonucleotide design to 7638. CAP3 has previously been shown to be tolerant of sequencing errors resulting from single pass sequencing and is effective at differentiating between closely related gene family members [14].

Putative sequence homology of each schistosome DNA element was assigned using the web-based Basic Local Alignment Search Tool (BLASTx) [15] searching against the NCBI protein non-redundant (nr) database. BLASTx hits with an Expect-value (*E*-value) of $\leq 10^{-05}$ were considered significant and the corresponding NCBI protein nr designation was used to annotate the EST contigs and singletons. BLASTx hits with an *E*-value of $\geq 10^{-05}$ were not considered significant and therefore the corresponding schistosome DNA elements obtaining these scores were annotated as 'UNKNOWN'. In a further attempt to annotate these unknown schistosome DNA sequences, they were compared against the *S. mansoni* EST database compiled by Verjovski-Almeida et al. [9] using BLASTn, where DNA elements generating BLASTn bit score values of ≥ 200 were considered a match. Subsequent annotation was provided then by the EST database assigned through additional BLASTx searches of NCBI (using their applied criteria [9]). Sequence similarity assignment was performed for two reasons: (1) to annotate the DNA sequence representations deposited on each DNA microarray; and (2) to identify which DNA strand to select for oligonucleotide design. For BLASTx searches with no significant similarity, the sense strand was selected for oligonucleotide design, unless there was a poly-T tract incorporated at one end of the parent DNA sequence (anti-sense strand was used in this case). Therefore, a small percentage of oligonucleotides deposited on this DNA microarray may have represented the non-coding DNA strand. The labelling procedure used in this study [16] (Klenow incorpo-