UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Construction of subtracted EST and normalised cDNA libraries from liver of chemical-exposed three-spined stickleback (Gasterosteus aculeatus) containing pollutant-responsive genes as a resource for transcriptome analysis

Brown, MM; Williams, Timothy; Chipman, James; Katsiadaki, I; Sanders, M; Craft, JA

DOI:
10.1016/j.marenvres.2008.02.043

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Brown, MM, Williams, T, Chipman, J, Katsiadaki, I, Sanders, M & Craft, JA 2008, 'Construction of subtracted EST and normalised cDNA libraries from liver of chemical-exposed three-spined stickleback (Gasterosteus aculeatus) containing pollutant-responsive genes as a resource for transcriptome analysis', *Marine Environmental Research*, vol. 66, no. 1, pp. 127-130. https://doi.org/10.1016/j.marenvres.2008.02.043

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 4/9/2015: Publishers version can be found at doi:10.1016/j.marenvres.2008.02.043

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 25. Apr. 2024

Accepted Manuscript

Construction of subtracted EST and normalised cDNA libraries from liver of chemical-exposed three-spined stickleback (*Gasterosteus aculeatus*) containing pollutant responsive genes as a resource for transcriptome analysis

Margaret M. Brown, Timothy D. Williams, J. Kevin Chipman, Ioanna Katsiadaki, Matthew Sanders, John A. Craft

PII: S0141-1136(08)00048-2
DOI: 10.1016/j.marenvres.2008.02.043
Reference: MERE 3200

To appear in: *Marine Environmental Research*

Please cite this article as: Brown, M.M., Williams, T.D., Chipman, J.K., Katsiadaki, I., Sanders, M., Craft, J.A., Construction of subtracted EST and normalised cDNA libraries from liver of chemical-exposed three-spined stickleback (*Gasterosteus aculeatus*) containing pollutant responsive genes as a resource for transcriptome analysis, *Marine Environmental Research* (2008), doi: 10.1016/j.marenvres.2008.02.043

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Construction of subtracted EST and normalised cDNA libraries from liver of chemical-exposed three-spined stickleback (*Gasterosteus aculeatus*) containing pollutant responsive genes as a resource for transcriptome analysis

Margaret M. Brown [a,*], **Timothy D. Williams** [b], **J. Kevin Chipman** [b], **Ioanna Katsiadaki** [c], **Matthew Sanders** [c], **John A. Craft** [a]

[a] *Department of Biological and Biomedical Sciences, Glasgow Caledonian University, UK;* [b] *School of Biosciences, The University of Birmingham, UK;* [c] *CEFAS, Weymouth laboratory, UK.*

Abstract

The three-spined stickleback (*Gasterosteus aculeatus*) is ideally suited to laboratory studies, while its wide distribution in the northern hemisphere gives it great potential as a sentinel organism. In the setting of a UK-wide collaboration (Fish Toxicogenomics) we have developed a microarray for transcriptomic analysis of chemical responses in populations of *G. aculeatus* under laboratory and field conditions. Although several EST libraries are available for this species none are from chemical-exposed fish and thus unlikely to include a full set of pollutant-responsive genes. To harvest such transcripts cDNA libraries were produced from liver of chemical-exposed mature males. Two normalised full-length libraries were generated

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

by different methods: 1) partial subtraction of polyA$^+$ RNA against solid-phase cDNA using magnetic bead technology; 2) degradation of double stranded cDNA formed by abundant transcripts. To enrich for pollutant-responsive genes a subtracted EST library was also generated. For each library ~1.5K clones were sequenced and characterised using Blast2GO. All libraries contained pollutant-responsive transcripts not previously available while additionally the subtracted library was generally enriched ~1.2 -10 fold for transcripts expected to be induced in response to the pollutants.

Keywords: Stickleback; cDNA; Subtractive suppression hybridisation; Microarray; normalisation

*Corresponding author: *m.a.brown@gcal.ac.uk* (M.A. Brown)

The three-spined stickleback (G*asterosteus aculeatus*) is a small euryhaline species with wide distribution. Its prevalence in freshwater systems in Northern Europe and suitability to laboratory conditions (short life cycle; easily cultivated; readily induced breeding cycle) make it highly informative in ecotoxicological studies, in particular of endocrine disruption (Wibe et al., 2002; Hahlbeck et al. 2004a; Hahlbeck et al., 2004b; Katsiadaki et al., 2006). Biology, behaviour and physiology have been studied extensively in *G. aculeatus* however, functional genomics knowledge is absent although as a major research organism for evolutionary biologists (Peichel, 2005) the genome is sequenced though not annotated. In the setting of a UK-wide collaboration (Fish Toxicogenomics -

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

www.biosciences.bham.ac.uk/fishtoxicogenomics) an expression microarray for *G. aculeatus* has been developed for analysis of chemical responses under lab and field conditions. To be useful in this context, an array must include pollutant-responsive cDNAs. Available EST libraries (Kingsley et al., 2004), derived from multiple tissues of non-exposed fish, are unlikely to contain full suites of pollutant-responsive liver transcripts. These transcripts must be chemically induced prior to generation of normalised libraries; or alternatively selected by subtraction of constitutively expressed transcripts from the transcriptome of chemical-exposed fish. To provide resources for our array both approaches were adopted to generate full-length cDNA and EST libraries.

To induce oestrogen-, heavy metal- and PAH-responsive genes, fish were exposed to ethynyloestradiol (EE2), copper (Cu) and di-benz(*a,h*)anthracene (DbA) respectively. For library construction polyA$^+$ RNA was isolated from liver of mature male fish exposed (24 or 48 hours) to either Cu, 10 or 100 µg/l; EE2, 10 or 100 ng/l; DbA, 1 or 10 µg/l or water. Full-length cDNA libraries were produced from pooled polyA$^+$ RNA (chemical-exposed and control samples (n=2) equal quantities from each time point/concentration) by alternate methods of normalisation: 1) partial-subtraction of polyA$^+$ RNA against solid-phase cDNA using magnetic bead technology (Ga_NmlG); 2) degradation of double stranded cDNA formed by abundant transcripts (Ga_NmlY). To remove abundant transcripts in Method 1, first-strand cDNA was synthesised directly on oligo(dT) cellulose magnetic beads, RNA template eluted and the solid-phase cDNA library hybridised with fresh pooled polyA$^+$ RNA. Resulting cDNA/RNA hybrids were collected by magnet and partially subtracted polyA$^+$ RNA remaining in solution subjected to two further subtractions prior to synthesis and linear amplification of cDNA (SMART™PCR cDNA Synthesis Kit, Clontech) (Zhu et

ACCEPTED MANUSCRIPT

al., 2001). This generated adequate material for construction of a phagemid library (SMART™ cDNA Library Construction Kit, Clontech). In Method 2, duplex-specific nuclease technology (DSN) (Zhulidov et al., 2004) was used for normalisation (DNA Trimmer Kit, Evrogen) and a plasmid library synthesised by long-distance PCR (Creator™ SMART™ cDNA Library Construction Kit, Clontech). For the subtracted EST library (Ga_S1TG), pooled polyA$^+$ RNA from liver of chemical-exposed fish (equal quantities from each time-point/concentration) was subtracted against pooled control (equal quantities from 24 and 48 hours) by suppression subtractive hybridisation (SSH) as described (Brown et al., 2004).

Randomly selected clones (~1.5K from each library) were subjected to one-pass sequencing and scored for quality using PHRED20, minimum 100 consecutive bases as criteria. Sequencing quality varied (Table 1), consistently good results were obtained with plasmid vector (Ga_NmlY) but 3' polyT structures within the phage vector (Ga_NmlG) and 3` polyA and the 5` GC rich region created by the SMART™ system within the ESTs (Ga_S1TG) caused a number to fail. All libraries contained pollutant-responsive transcripts not previously available in stickleback, however, the subtracted library was generally enriched ~1.2-10-fold for transcripts expected to be induced by the model chemicals.

Of the readable sequences 1817/3568 did not form contigs with those from the Kingsley library; of these 833 were singlets the remainder forming contigs within the libraries. Taking into account redundancy levels (Table 1) about 49% of the1817 sequences represent unique clones not previously available however, several genes may be represented although not contiguous. It was not expected to infer which genes were altered specific to the pollutants, this will be discerned downstream by back-screening of the samples against the arrayed libraries.

ACCEPTED MANUSCRIPT

Gene Ontology based data mining (Blast2GO (Conesa et al., 2005)) showed the libraries covered a wide range of biological functions. A small number of transcripts were universally over-represented, mainly ribosomal or those involved in protein synthesis, but generally gene identity profiles differed between libraries. In the normalised libraries this could be due to the relatively small numbers sequenced however it is noteworthy there was less redundancy in that normalised by partial subtraction (Method 1). The subtracted library was more enriched for chemical-responsive transcripts both known (e.g.s CYP1A, metallothionein, vitellogenin) and novel to this species (e.g.s DbA: fatty acid binding proteins, EE2: insulin-like growth factor binding protein; Cu: betaine homocysteine methyltransferase) (Table 2). Blast2GO, a powerful annotation tool for high-throughput processing of sequence data, has been shown to reach 60-70% accuracy with datasets where annotation/functionality is known. With ESTs from non-model species lower accuracy would be expected and more detailed manual annotation of predicted full length cDNAs extracted from the genomic database (http://www.ensembl.org) would be desirable but labour intensive.

Provision of resources for expression arrays must ideally be comprehensive for detection of gene events specific to chemical exposure. The methods used here for library construction sought to maximise transcriptome coverage accordingly. Starting material was pooled to produce single libraries containing transcripts responsive to several pollutants, an economy on time and resources. The resulting libraries were complementary and provided coverage of the transcriptome with the subtracted EST library more highly enriched for pollutant-responsive transcripts. The combined use of complementary cDNA libraries is therefore advantageous to isolate tissue or treatment specific transcripts.

References

Brown, M., Robinson, C., Davies, I.M., Moffat, C.F., Redshaw, J., and Craft, J.A. (2004). Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 552, 35-49.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005). Bioinformatics, 21, 3674-3676.

Hahlbeck, E., Griffiths, R., and Bengtsson, B.E. (2004a). Aquatic Toxicology, 70, 287-310.

Hahlbeck, E., Katsiadaki, I., Mayer, I., Adolfsson-Erici, M., James, J., and Bengtsson, B.-E. (2004b). Aquatic Toxicology, 70,311-326.

Katsiadaki, I., Morris, S., Squires, C., Hurst, M.R., James, J.D., and Scott, A.P. (2006). Environmental Health Perspectives 114, 115-121.

Kingsley, D.M., Zhu, B., Osoegawa, K., de Jong, P.J., Schein, J., Marra, M., et al. (2004). Behaviour 141, 1331-1344.

Peichel, C.L. (2005). Developmental Dynamics, 234, 815-823.

Wibe, A.E., Rosenqvist, G., and Jenssen, B.M. (2002). Environmental Research, 90, 136-141.

Zhu, Y.Y., Machleder, E.M., Chenchik, A.C., Li, R., and Siebert, P.D. (2001). Biotechniques, 30, 892-897.

Zhulidov, P.A., Bogdanova, E.A., Shcheglov, A.S., Vagner, L.L., Khaspekov, G.L., Kozhemyako, V.B., et al. (2004). Nucleic Acid Research 32, e37,1-8.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1.

Comparison of sequencing quality and Blast2GO annotation of the normalised and subtracted libraries.

LIBRARY	Ga_NmlG (directional cloning) Norm/Method 1	Ga_NmlY (directional cloning) Norm/Method 2	Ga_S1TG (non-directional cloning) Subtracted SSH
QUALITY SEQUENCE	751/1191 (63%)	1275/1440 (89%)	1542/2016 (77%)
NO HOMOLOG	127 (17%)	69 (5%)	73 (23%)
SIMILAR TO UNKNOWN	232 (31%)	221 (17%)	686 (45%)
ANNOTATED TRANSCRIPTS	392 (52%)	985 (78%)	783 (32%)
UNIQUE IDS	262/392 (67%)	394/985 (40%)	323/783 (40%)
GO ASSOCIATIONS (BIOLOGICAL PROCESS/ MOLECULAR FUNCTION)	121	182	177

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2.

ESTs from transcripts up-regulated in response to chemical exposures (EE2, dibenzoanthracene, copper) were isolated from liver of *G. aculeatus* by suppression subtractive hybridisation (treated minus control). Randomly selected clones were subjected to one-pass sequencing and annotated using Blast2GO. All sequences are available at NCBI, dbEST accession numbers EG588073 – EG591696.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Process	Accession	Most similar to
Phase I metabolism	EG590952	cytochrome P450 1A
	EG591310	Cytochrome P450 2K1 (CYPIIK1) (P450 LMC2)
	EG590829	cytochrome P450 2K5
	EG591338	cytochrome P450 2N1
	EG590214	cytochrome P450 2P2
	EG590878	cytochrome P450 3A69
	EG590659	cytochrome P450 monooxygenase CYP2K6
	EG590599	cytochrome P450, family 2, subfamily J, polypeptide 2, B (CYP2J2B)
	EG590223	cytochrome P450, family 46, subfamily a, polypeptide 1 (CYP46A1)
	EG591178	cytochrome P450, (CYP2J) (arachidonic acid epoxygenase)
	EG590478	similar to Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYPVII) (Cholesterol 7-alpha-hydroxylase)
	EG591390	similar to family 4 cytochrome P450 isoform 1
Phase II metabolism	EG588837	glutathione S-transferase
	EG590578	PREDICTED: similar to UDP-glucuronosyltransferase 2A1 precursor, microsomal
	EG590674	similar to sulfotransferase family 3A, member 1
	EG591152	UDP-glucuronosyltransferase (UDPGT) UGT3
multidrug resistance	EG590677	PREDICTED: similar to solute carrier family 22 (organic anion transporter), member 20, partial
estrogen responsive	EG590483	14kDa apolipoprotein
	EG590549	17-beta hydroxysteroid dehydrogenase type 3
	EG590466	3-oxo-5 alpha-steroid 4-dehydrogenase 2
	EG590205	B-cadherin
	EG591685	beta globin mRNA, complete cds
	EG590420	beta tubulin
	EG591549	beta-2 microglobulin precursor
	EG591437	cadherin 1, epithelial
	EG590927	chitinase3
	EG590252	cholesteryl ester transfer protein
	EG591502	choriogenin H
	EG590704	choriogenin L
	EG591274	chorion protein
	EG590432	diazepam binding inhibitor (GABA receptor modulator, acyl-Coenzyme A binding protein)
	EG590598	drtp1
	EG590361	similar to insulin-like growth factor 2 precursor (LOC572929), mRNA
	EG590963	VgC mRNA for phosvitinless vitellogenin, complete cds
	EG590342	vitellogenin
	EG591213	zona pellucida protein
	EG590468	PREDICTED: similar to calmodulin
oxidative stress	EG591481	15 kDa selenoprotein precursor
	EG590673	acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase)
	EG590254	acyl-CoA thioesterase 1
	EG590993	adenine nucleotide translocator
	EG590934	alanine-glyoxylate aminotransferase
	EG590172	aldolase B
	EG591220	Annexin A5
	EG590467	bile acid-Coenzyme A: amino acid N-acyltransferase
	EG591607	clone C232 ubiquitin mRNA, partial cds
	EG591484	clusterin-2 protein
	EG590882	glutathione peroxidase
	EG590546	peroxiredoxin 4
	EG591402	peroxisomal acyl-CoA thioesterase 2A
	EG590575	phospholipid hydroperoxide glutathione peroxidase A
	EG591687	PREDICTED: similar to NADH-ubiquinone oxidoreductase 9 kDa subunit, mitochondrial precursor (Complex I-9KD) (CI-9KD)
	EG591340	PREDICTED: similar to ubiquitin isoform 1
	EG590909	similar to 5-lipoxygenase activating protein (FLAP) (MK-886-binding protein)
	EG590471	ubiquitin-binding protein homolog
	EG590169	ubiquitin-conjugating enzyme E2N-like
	EG591131	thioredoxin-like 4A (txnl4a), mRNA
copper responsive proteins	EG590419	COMM domain containing 3
	EG588881	betaine homocysteine methyltransferase
	EG590588	cytochrome c oxidase subunit I
	EG591098	cytochrome c oxidase subunit II
	EG590623	cytochrome c oxidase subunit III
	EG590661	cytochrome c oxidase, subunit VIIc
	EG591694	PREDICTED: similar to retinol binding protein 7, cellular
	EG588214	Metallothionein
fatty acid metabolism	EG591416	fatty acid binding protein H8-isoform
	EG590708	fatty acid-binding protein
	EG591611	Hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase
	EG590737	liver-basic fatty acid binding protein
	EG590662	acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase)
	EG590299	TBT-binding protein mRNA, partial cds
cholesterol metabolism	EG590393	transmembrane 7 superfamily member 2
	EG590455	isopentenyl-diphosphate delta isomerase
	EG591273	farnesyl diphosphate synthase (dimethylallyltranstransferase, geranyltranstransferase)
	EG590478	similar to Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYPVII) (Cholesterol 7-alpha-hydroxylase)
	EG590252	cholesteryl ester transfer protein
transcriptional repression	EG591387	RNA-binding protein VgRBP71
apoptosis	EG591110	PREDICTED: similar to Ras GTPase-activating-like protein IQGAP1 (P195)
	EG591612	cell death-regulatory protein GRIM19
	EG590908	PEST-containing nuclear protein
	EG590888	metaxin 2

ACCEPTED MANUSCRIPT