NEW MOLECULAR TOOLS: APPLICATION OF THE µAQUA PHYLOCHIP AND CONCOMITANT FISH PROBES TO STUDY FRESHWATER PATHOGENS FROM SAMPLES TAKEN ALONG THE TIBER RIVER, ITALY

L. K. MEDLIN[1,2], D. GUILLEBAULT[2], G. MENGS[3], C. GARBI[3], L. DEJANA[3], C. FAJARDO[3] & M. MARTIN[4]

[1]Marine Biological Association of the UK, UK
[2]Microbia Environnement, France
[3]Natural Biotec, Spain
[4]University Computense of Madrid, Spain

ABSTRACT

Current knowledge about aquatic pathogens are scarce because bacteria, protozoans, algae and their toxins occur at low concentrations, making them difficult to measure directly or to filter sufficient volumes to facilitate detection. We developed and validated tools to detect pathogens in freshwater systems. To evaluate impacts on water quality, a phylogenetic microarray was developed in the EU project µAQUA to detect simultaneously numerous pathogens and was applied in MicroCoKit, to samples taken from four locations from two seasons for two years along the length of the Tiber River, Italy. The sites ranged from a pristine site near its source to ones contaminated by agricultural, industrial and anthropogenic waste moving downstream to near its mouth. Fifty litres were collected and concentrated using a hollow-fibre ultrafiltration, a rapid method with minimal cell loss to provide a concentrate for downstream analysis. The 60 Da cut-off ensures many organics, such as toxins, will be concentrated for analysis. Aliquots from the concentrate were preserved in TRI-Reagent and total RNA extracted, labelled and hybridised to the phylochip to detect pathogenic bacteria, protozoa and toxic cyanobacteria. The microarray results gave positive signals for all pathogens. Calibration curves enabled us to infer cell concentrations. Cross validation was performed using FISH probes for selected toxic cyanobacteria and hybridised to aliquots taken from the raw water prior to concentration by the hollow fibre filters.

Keywords: bacteria, cyanobacteria, hollow fibre filters, phylochips, monitoring, pathogens, protozoa.

1 INTRODUCTION

Monitoring drinking water quality is an important public health issue. The EU Project µAQUA made novel tools for the early and sensitive detection of water-borne pathogens (bacteria, cyanobacteria and protozoans) and toxins. Pathogenic organisms occurring in lakes and rivers used as drinking water reservoirs represent a serious health-hazard. River water is usually contaminated with bacteria (*E. coli, C. perfringens,* etc.), viruses (adenoviruses) and pathogenic protozoa (*G. duodenalis, C. parvum,* etc.). Cyanobacteria are more of a problem in lakes where they can bloom.

Any species can be detected through the use of DNA/RNA probes (barcodes) used in a microarray detection platform, which consists of the barcodes applied to the surface of a specially treated glass slides in a dot blot fashion [1], [2]. Microarrays, originally used in functional genomics for studying gene expression, etc. (see review in [3]), quickly expanded into species identification, the phylochip [4]. Phylochips can replace traditional methods,

which are laborious, technically demanding and time-consuming, and are faster, more accurate and can be quantified for monitoring purposes [5], [6].

The µAQUA microarray was field tested in six European countries from rivers to lakes to tap water over two years. Validation was performed by four µAQUA partners using either traditional cell counting regimes (cyanobacterial) to plate counts (bacteria) to magnetic beds/antibodies (protozoa) with good correlations between both methods. Results for three of the monitoring sites are published (the Tiber River above and below Rome, Italy [7], rivers near Paris, France [8], one lake in Bulgaria (9]). The µAQUA microarray was transferred into the EU project, MICROCOKIT (www.microcokit.eu), who sampled the entire length of the Tiber River in spring and autumn for two years, concentrating on four sites with pollution distinct types: pristine, agricultural, industrial and anthropogenic. One site was near the Rome sampling site from µAQUA. In MICROCOKIT, next generation sequencing, quantitative PCR, and Fluorescent *in-situ* Hybridisation (FISH) methods will validate the microarray. We present here microarray results and their validation with FISH.

2 MATERIAL AND METHODS

2.1 Sampling regime

Four sites along the length of the Tiber River were identified as sources of different types of pollution (Fig. 1). Samples were collected and processed within one day, once in the spring and in the autumn of two consecutive years. Fifty litres of river water were filtered through a polysulfone hollow-fibre ultrafilter module (HF80S, Fresenius Medical Care, Bad Homberg, Germany) by the MICROCOKIT team (JRC and CNR-IRSA). Filtration and cell concentration process is described in detail in [7]–[9]. Thirty mls of the concentrate were preserved in Tri-reagent (Sigma, France) at a ratio of 1:4 concentrate:Tri-reagent to ensure sufficient concentrate dilution for optimal RNA extraction [3], [10], [11].

2.2 Probe design

Protocols used in µAQUA were adopted from the MIDTAL project for toxic algae [10]. Probes for species, genera, classes or phyla of pathogenic bacteria, toxic cyanobacteria, pathogenic protozoa and diatoms as indicator species of water quality were either collected from the literature and extended to 25 nts or newly designed following a hierarchical fashion. Thus, for a species or genus to be present, the taxonomic hierarchy leading to that taxon had to be present (Table 1), to prevent false positives. All probes were checked *in silico* and their biophysical properties analyzed. Positive control probes and higher taxonomic probes targeting kingdom and phylum levels came from MIDTAL [11], Both arrays are available from Microbia Environnement (Banyuls/Mer, France).

Probes, whose results are presented here, were designed for these bacterial genera and species: *Salmonella spp., Shigella spp., Campylobacter, C. coli and C. jejuni, E. coli, Legionella, L. pneumophila, Clostridium perfringens, Listeria, L. monocytogenes, Staphylococcus aureus, Yersinia, Y. enterocolitica, Vibrionaceae*, *Aeromonas*, *Bacillus cereus, Pseudomonas*, and *Mycobacterium*; for these cyanobacterial species and genera: *Microcystis aeruginosa, Planktothrix, P. agardhii, Nodularia spumigena, Anabaena., Aphanizomenon flos-aquae, Cylindrospermopsis*; and for these protozoans: *Cryptosporidium, Giardia, Entamoeba, Naegleria;* for higher group level probes were designed ranging from family, order, class, phylum to domain depending on target availability (hierarchy in Table 1)

Table 1: Summary of hierarchical probes for Cyanobacteria, Bacteria and Protozoa. (Read the hierarchy from left to right beginning with either a species or a genus level probe.)

A. Hierarchy for Cyanobacteria

Species	**Genus or multiple genera**	**Order**	**2nd level of multiple orders[1]**	**3rd level of multiple orders[1]**	**Phylum/division**	**Kingdom**
	Microcystis	Chlorococcales	Chroococales/ Nostocales		Cyanobacteria	Eubacteria
	Anabaena/ Aphanizomenon/ Cylindrospermopsis	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Anabaena/ Aphanizomenon	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Anabaena/ Aphanizomenon/ Nodularia	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Aphanizomenon	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Cylindrospermopsis	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Nodularia/ Anabaena	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Nodularia/Anabaena/ Aphanizomenon	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
	Nodularia	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
Nodularia spumigena	*Nodularia*	Nostocales	Oscillatoriales/ Nostocales	Chroococales/ Nostocales	Cyanobacteria	Eubacteria
Planktothrix agardhii	*Planktothrix*	Oscillatoriales	Nostocales/ Oscillatoriales		Cyanobacteria	Eubacteria

[1]The second name is the stronger probe and the probe in the hierarchy.

Table 1: Continued.

B. Hierarchy for Bacteria

Species	Genus	Family/class	Phylum/division	Kingdom
	E. coli/Shigella	Enterobacteriaceae	Gammaproteobacteria	Eubacteria
	E. coli/Shigella/Salmonella	Enterobacteriaceae	Gammaproteobacteria	Eubacteria
	Salmonella	Enterobacteriaceae	Gammaproteobacteria	Eubacteria
Yersinia enterocolotica	*Yersinia*	Enterobacteriaceae	Gammaproteobacteria	Eubacteria
	Vibrio		Gammaproteobacteria	Eubacteria
Legionella pneumophila	*Legionella*		Gammaproteobacteria	Eubacteria
	Pseudomonas		Gammaproteobacteria	Eubacteria
Campylobacter coli/jejuni	*Campylobacter*			Eubacteria
Staphylococcus aureus	*Staphylococcus*	Bacilli	Firmicutes	Eubacteria
	Listeria	Bacilli	Firmicutes	Eubacteria
	Bacillus spp	Bacilli	Firmicutes	Eubacteria
Clostridium perfringens			Firmicutes	Eubacteria
	Mycobacterium		Actinobacteria	Eubacteria

C. Hierarchy for Protozoa

Species	Genus	Family	Order	Class	Kingdom
Cryptosporidium hominis	*Cryptosporidium*	Cryptosporidae	Eimeriorina	Conoidasida	Eukaryote
Cryptosporidium parvum/hominis					
Cryptosporidium parvum/hominis					
Cryptosporidium parvum					
Giardia intestinalis	*Giardia*			Diplomonadida	Eukaryote
Naegleria fowleri	*Naegleria*				Eukaryote
Naegleria lovaniensis					
Entamoeba histolytica	*Entamoeba*				Eukaryote

Figure 1: Map of the Tiber River showing the four sampling sites. T1 is a pristine site located near the source of the river. T2 is at Attigliano, with agricultural pollution; T3 where the Aniene River joins the Tiber near industrial pollution; T4 at Scafa and is subjected to anthropogenic pollution.

2.3 Microarray construction, hybridization and analysis

246 probes were spotted by Scienion AG (Berlin, Germany) as follows: each microarray slide contained two arrays with eight replicates for each probe. Hybridisation of each sample was performed on 2 different slides, thus producing a pseudo-replicate of 16 spots. Hybridizations were performed according to [3] and described in detail in [6]. Hybridization results were analyzed using the hierarchy file designed with the Phylochip analyzer program [12] and the GPR-Analyzer v1.28 [13]. Signals were eliminated if the hierarchy was broken. We generated a heatmap for the microarray signal for each probe using PermutMatrix [14] (http://www.atgc-montpellier.fr/permutmatrix/).

2.4 Calibration curves for cyanobacteria quantification using microarray

Cyanobacteria were quantified by cell counts using the Utermöhl method. RNA was extracted from known cell numbers from pure cultures of *Microcystis*, *Planktothrix*, *Nodularia, Aphanizomenon* and *Anabaena.* Microarray analyses were performed with RNA equivalent to 10,000, 50,000, 100,000 and 500,000 cells for concentration curves to convert signals to cell numbers/L (Table 2).

2.5 FISH probe design and hybridisation

Microarray probes for *Microcystis* + *M. aeruginosa* and *Planktothrix* + *P. agardhii* were shortened to match melting temperatures for genus + species probes to be used at the same time and for ease of cell penetration (Table 3). The probes were labelled with either FITC or Cy3. The idea of using a CY3 label on the genus level probe was to use them in a hierarchical fashion: once a green FITC species-specific signal was detected, then filters were switched to the CY3 channel to see if the CY3-labeled genus was there. Different fixation methods were tried to minimise the autofluorescence of the chlorophyll. 200 µl of pure cultures of *M. aeruginosa* and *P. agardhii* were mixed with 800 µl PBS (1.10^2 cell/ml), vortexed and filtered under vacuum onto a polycarbonate 0,2 µm GTTP Millipore filter. One hour and an overnight of fixation time with Saline-ETOH (see recipes in [15]) were tried with and without one hour of 50% dimethylformamide (DMF) [16] to bleach further the chlorophyll auto fluorescence. Optimal fixation method was Saline-ETOH 1 hour + one hour DMF. Two formamide concentrations (15% and 20%) with increasing temperatures (45ºC, 47ºC, 49ºC

Table 2: Cell numbers inferred from empirically determined calibration curves from pure cultures.

Probes	T2 Mar 14	T3 Mar 14	T4 Mar 14	T1 Oct 14	T2 Oct 14	T3 Oct 14	T4 Oct 14	T1 Mar 15	T2 Mar 15	T3 Mar 15	T4 Mar 15	T1 Oct 15	T2 Oct 15	T3 Oct 15	T4 Oct 15
GNMicrS01	10	80	0	0	2219	0	0	0	8	466	5	5	5	5	5
GnMicrS03	0	0	0	0	1973	4206	0	0	1	598	1	1	1	1	1
GNMicS02	0	16	0	0	2042	3096	0	34	14	302	2	2	2	2	2
GNMicS04	291	329	0	0	20834	10705	7709	456	85	1391	2	2	2	2	2
Average Genus Level	75	106	0	0	6767	4502	1927	122	27	689	3	3	3	3	3
GNCylS01	86	100	199	22	563	3496	680	30	217	129	47	47	47	47	47
GNPlankS02	275	594	0	472	3423	9372	5544	290	789	1516	104	104	104	104	104
GNAphPlaS01=PLKTX	9	0	336	58	12699	2297	2590	6	27	242	1	1	1	1	1
Average Genus Level	142	297	168	265	8061	5834	4067	148	408	879	53	53	53	53	53
GNNodS01	7	26	51	0	224	933	439	6	33	106	22	22	22	22	22
GNNodS02	0	0	0	0	0	13014	0	244	0	1201	1	1	1	1	1
Average Genus Level	4	13	25	0	112	6973	220	125	16	654	12	12	12	12	12
SPNodSpuS02	0	0	0	89	662	2009	0	18	140	338	2	2	2	2	2
GNAphS01	0	0	0	0	0	5952	0	112	0	549	1	1	1	1	1
GNAnaAphS01=ANA major	1	19	34	36	645	608	355	18	0	112	2	2	2	2	2
GNAnaAphS02=ANA major	0	9	0	0	2285	0	756	0	5	145	1	1	1	1	1
GNAnaAphS03=ANA major	1	0	0	50	1270	547	893	0	6	103	0	0	0	0	0
GNAphAnaS01=ANA major	0	0	0	42	443	0	0	0	7	104	1	1	1	1	1
Average Anabaena clade	0	7	9	32	1161	289	501	5	5	116	1	1	1	1	1
GNAphAnaS02=APHA major	385	59	161	0	14963	6175	23783	0	1150	1123	169	169	169	169	169
GNAphAnaS04=APHA major	0	0	0	0	3816	3353	0	107	15	560	1	1	1	1	1
Average Aphanizomenon clade	96	15	40	0	5535	3870	5946	55	292	682	43	43	43	43	43

Table 3: RNA probes for Cyanobacterial FISH hybridisation.

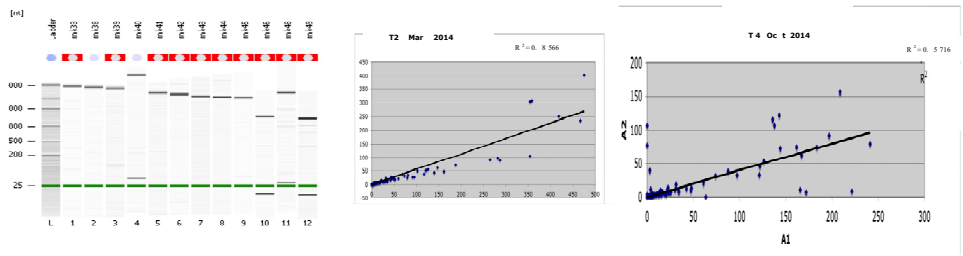
Probe name	Species	Gene	Sequences 5'-3'- fluorochrome
GNPlankS02	*Planktothrix spp.*	16S rRNA	TCAAGGAGATTCCAGAGATGTCAAGT-CY3
PkAgD03	*Planktothrix agardhii*	23S rRNA	CTCTCTAAGTCCAGTGTCGCTG-FITC
GNMICSO5	*Microcystis spp.*	16S rRNA	GCGTGAGGGAGGAAGGTCTTT-CY3
MicAerD03	*Microcystis aeruginosa*	23S rRNA	CTTGATCAACGCCACTTCCCTC FITC

and 51ºC) were performed to optimise the FISH method. The best hybridisation conditions were 20% formamide at 49ºC. *E. coli* OP50 was the negative control and pure cultures of each alga, the positive controls. Two mls of raw Tiber River water from each sampling site was filtered onto a polycarbonate 0.2 µm GTTP Millipore filter, which was cut into sectors for hybridisation and then mounted with Prolongol Mounting Medium for examination and counting with an Olympus Laser Confocal Microscope at the UCM Microscopy Service.

3 RESULTS AND DISCUSSION

3.1 Microarray analysis

RNA was successfully extracted using MIDTAL and µAQUA RNA extraction protocols. RNA quality was sufficiently high to proceed to hybridisation (Fig. 2(A)). A low amount of RNA was extracted in March 2014 and its hybridisation signals were below background and unusable. Total intensity from each replicate was compared in a regression analysis to determine if the two hybridisations are true replicates with no need to repeat hybridisations (Table 5). Fig. 2(B), (C) show an example of the good and a poorer regression. The lower reproducibility of October 2014 T4 and October 2015 T3 as compared to the other samples was because one showed a stronger signal for some probes than the other, which is likely a spotting error. They were not redone because no probes were missing between the two and were averaged.



(A) (B) (C)

Figure 2: (A) Bioanalyser gel showing typical RNA quality from the environmental samples. (B), (C) Regression of the two replicate microarray grids, (B) shows good replication, (C) shows poorer replication; no missing probes, only large differences in intensities.

Table 4: RNA probes for Cyanobacterial FISH hybridisation.

Probe name	Species	Gene	Sequences 5'-3'- fluorochrome
GNPlankS02	*Planktothrix spp.*	16S rRNA	TCAAGGAGATTCCAGAGATGTCAAGT-CY3
PkAgD03	*Planktothrix agardhii*	23S rRNA	CTCTCTAAGTCCAGTGTCGCTG-FITC
GNMICSO5	*Microcystis spp.*	16S rRNA	GCGTGAGGGAGGAAGGTCTTT-CY3
MicAerD03	*Microcystis aeruginosa*	23S rRNA	CTTGATCAACGCCACTTCCCTC FITC

Table 5: Comparison of the regression (R^2) of the two pseudo-replicate hybridisations for each sampling date.

T1 Not enough RNA		T2Mar 2014	0.86	T3 Mar 2014	0.99	T4 Mar 2014	0.99
T1 Oct 2014	0.88	T2 Oct 2014	0.91	T3 Oct 2014	0.96	T4 Oct 2014	0.57
T1 Mar 2015	0.76	T2 Mar 2015	0.97	T3 Mar 2015	0.87	T4 Mar 2015	0.93
T1 Oct 2015	0.78	T2 Oct 2015	0.93	T3 Oct 2015	0.77	T4 Oct 2015	0.81

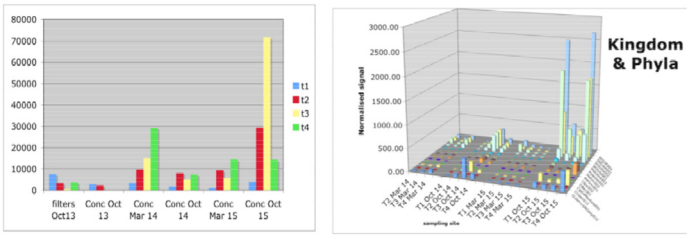
3.2 Total community

Our microarray successfully detected target pathogens in the Tiber River. The sampling sites appeared to be unique, with T2, the agriculturally polluted site and T4, the site receiving anthropogenic impact, likely responded to nutrient loading, which increased the presence of the pathogens at those sites. Throughout the Tiber, bacterial load was high with some evidence of seasonality: higher signals in the autumn than in the spring, when more rain and runoff occurs. Total community (both eukaryotes and prokaryotes) hybridising to the microarray were more abundant by 3–4 times in October 2015 than any other time, and this was reflected in the total RNA extracted (Fig. 4(A)). In October 2015, T4 had the highest bacterial and eukaryote signal (Fig. 4(B)). Marcheggiani et al. [7] also sampling the Tiber River with bacterial plate count confirmation, also found higher signals in the autumn, after the dry summer season.

3.3 Bacteria

3.3.1 Kingdom and Phylum level (Fig. 3)

March 2014. At the Kingdom level, there were stronger Eubacterial signals at site T2 and T4 than at T3. Of the bacterial phyla on the microarray, Actinobacteria, Firmicutes, and Gamma proteobacteria were present at all three sites, with the highest signals being obtained at T4. *Bacillus spp*. signals were not recorded at T2 or T4, with a small signal at T3. *October 2014*. At T1, there was only a slight signal in the Eubacterial probes. In contrast to the spring sample, T3 had the highest Eubacterial signal, most of which could be attributed to Gammaproteobacteria, although other classes were also present. *March 2015*. The lower signal of the EUB probes at T3 and T4 in March 2015 reflects either that many of the bacteria were not recognised by this probe or that there is sufficient overlap in the regions of the



(A) (B)

Figure 3: (A) Total RNA extracted over two-year sampling period. October 2013 represents a practice sampling and was not used in any hybridisations. (B) Microarray signals at the Kingdom and Phylum level.

probes that probes for other species and genera were more easily hybridised to the array leaving less RNA available to the EUB probes. *October 2015.* Signals were extremely high at T4 with Eubacterial signals being six times more than other times; most were Gammaproteobacteria. Class Bacillus was poorly represented. However T1, purportedly a pristine site had an extremely high bacterial and eukaryotic load. Target pathogens were present and confirmed by hierarchy tests (Figs 4–6).

3.3.2 Hierarchical probes for *Escherichia*, *Salmonella*, *Yersinia*, *Listeria*, *Pseudomonas* and *Vibrio* (Fig. 4)

March 2014. Some. Family and genus level probes recognising *Escherichia* were recorded at T4. *Listeria* is recorded by three genus level probes at sites T3 and T4. *Salmonella* and *Yersinia* are recorded by one genus probe at T4 and at T3 and T4, respectively. *Vibrio* was not present. *Pseudomonas* was strongly recorded by 4 probes at T2, T3 and T4. *October 2014.* Many target pathogens were present. Family/genus level probes for *Escherichia* were recorded at T2, T3, and T4. *Salmonella* and *Yersinia* are recorded at all three sites and *Y. enterocolytica* is recorded by two species level probes at all three sites. *Pseudomonas* was recorded by 4 genus probes at T2, T3 and T4; *Vibrio* is present with one genus probe at all three sites. *Legionella* is present at all three sites and *L. pneumophila* is present at T3 and T4. *March 2015.* T1 continued to be pristine with no pathogens recorded. *Escherichia, Salmonella, Listeria,* and *Vibrio* and *Y. enterolytica* were present at T3. *Pseudomonas* was recorded at T2, T3, and T4. *October 2015.* T1 appeared to have received some kind of contamination because *E. coli,* and *Pseudomonas* were present. *Y. enterocolytica* is recorded at T2, 3, 4. *E. coli* was recorded at T3 and 4; *Listeria* was present at T4.

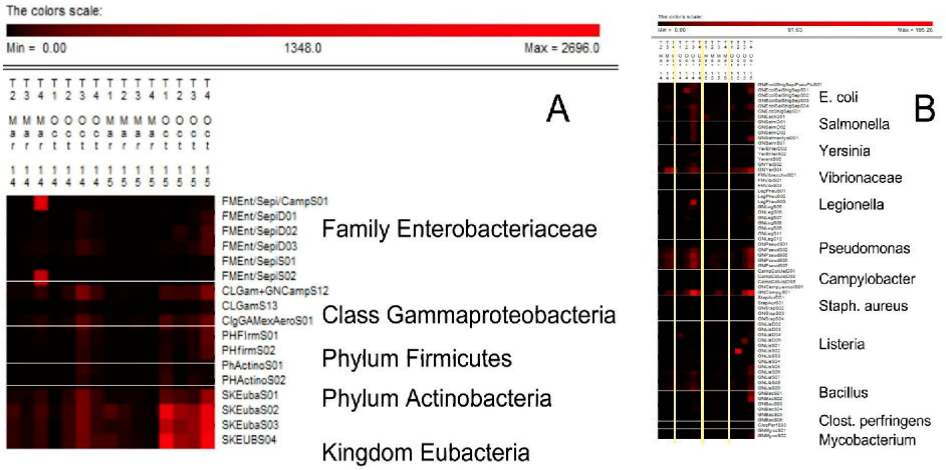


Figure 4: Heatmap of the relative abundance of (A) the bacterial hierarchical probes from family to kingdom and (B) from genus to species. Vertical lines separate sampling dates.

3.3.3 Hierarchical probes for *Legionella*, *Bacillus*, *Clostridium*, *Mycobacterium*, *Campylobacter*, and *Staphylococcus* (Fig. 4)

March 2014. No target bacteria from this group were present during this sampling except for one genus level probe for *Legionella* at T2 and T4. *October 2014*. *Mycobacterium*, *Bacillus* and *Campylobacter* are recorded at T3. The *C. jejuni* signal is likely a false positive because genus level probes were not present. *Staphylococcus* is recorded rarely but S. aureus is not recorded. *Legionella* is recorded by 6 genus probes at three sites and at the species level at T3 and T4. *Clostridium* is absent from all sites. *March 2015*. All target bacteria at the genus level and the species *L. pneumophila*, *C. perfringens*, *S. aureus* and *C. coli/jejuni* are recorded at T3. *October 2015*. All target bacteria at the genus level and *C. perfringens* and *C. coli/jejuni* are recorded at T4.

3.4 Cyanobacteria

3.4.1 Kingdom, Phyla and order level (Fig. 5)

Cyanobacteria were present though out the river except at T1 and were primarily represented by filamentous forms of the order Oscillatoriales. Differentiation among the different sites and seasons was more easily seen at the genus/species level, where T2, with the agricultural pollution, had the highest signals and more unicellular species. Phylum probes reacted, with some, noticeable stronger than others. All but three of the order level probes produced a signal in March 2014. In October 2014, the cyanobacterial community was highly diverse with all order level probes producing a signal even some present at T1. Cyanobacteria were poorly represented in March 2015. In October 2015, signals were the highest achieved, with most biomass in Chroococales and Oscillatoriales. 2014 was more diverse than in 2015.

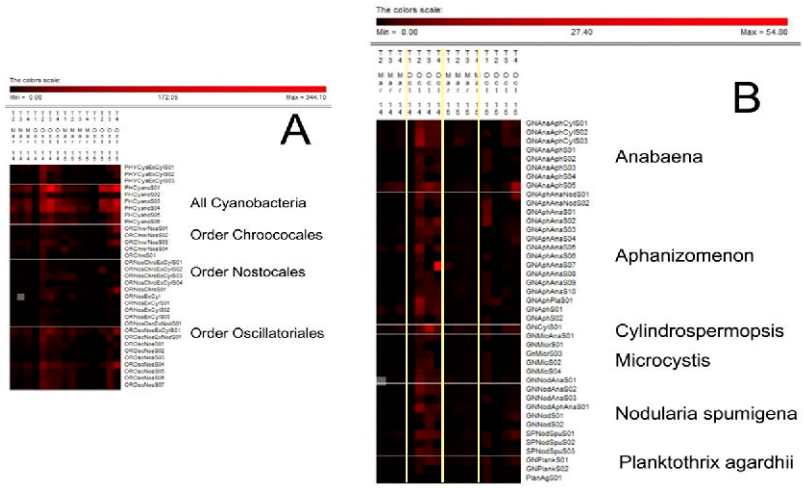


Figure 5: Heatmap of the relative abundance of (A) cyanobacterial order to phylum and (B) genus to species hierarchical probes. Vertical lines separate sampling dates.

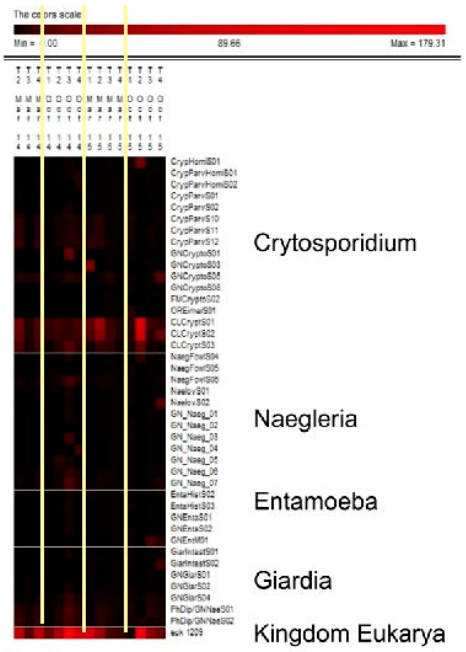


Figure 6: Heatmap showing the relative abundance of the protozoan hierarchical probes from the species level to Kingdom. Vertical lines separate sampling dates.

3.4.2 Toxic cyanobacterial genera and species probes (Fig. 5)

March 2014. Planktothrix, Microcystis, Cylindrospermopsis, Aphanizomenon/ Anabaena and *Nodularia* were present at all three sites. *N. spumigena* and *P. agardhii* were present at all three sites and at site T2, respectively. *October 2014.* All toxic genera and species produced a strong signal at T2, 3, and 4. *March 2015. Microcystis, Cylindrospermopsis, Nodularia,* and *Planktothrix* were present at T3. *October 2015. Microcystis, Cylindrospermopsis, Nodularia,* and *Planktothrix* were present at all except T2. With our calibration curves for the cyanobacteria, we can infer cell numbers from the signals (Table 3). Each probe has its own strength so a slightly different cell number will be inferred from the probes for a given species and averages for each species are made.

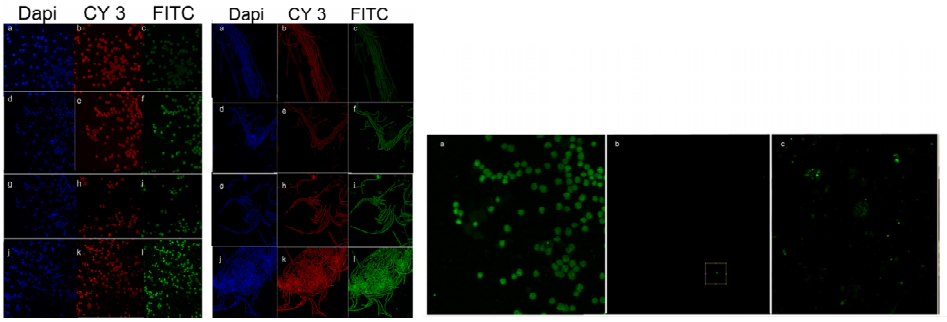
3.4.3 Protozoans (Fig. 6)

Higher group probes for intestinal protozoan parasites (*Giardia, Naegleria, Entamoeba* and *Cryptosporidium)* showed strong responses throughout the sampling period. Other genera/species, which may or may not infect humans, may also be present along with our target species. *March 2014. C. parvum* (complete hierarchy except for the order and family level) was recorded at T2 and T4, whereas the genus and class level probes were highlighted at all three sites indicating others present. *N. fowleri* was highlighted at T2 and T4. Phylum level probes for *Giardia* were highlighted, suggesting other genera may be present. *October 2014*. *C. parvum* and *hominis* (complete hierarchy) were present at all four sites. *Entamoeba*, *Giardia* and *Naegleria* were present at the genus and species level at T3. *March 2015*. *Cryptosporidium parvum* (complete hierarchy) was present at T1 and T2. *Naegleria* is

present at all four sites. *Entamoeba* and *Giardia* were present at T3. October 2015. Class Conoidasida was present at all sites, particularly at T1 and 2 but *C. parvum* was not present because the genus level probes did not react except for a low signal at T1. *N. flowleri* was present at T1, 3 and 4, whereas *N. lovenensis* was present at T1 and 4. *G. intestinalis* and *Entamoeba* were present at T3 and 4.

3.5 FISH Analysis (Fig. 7)

FITC probes for *M. aeruginosa* and *P. agardhii* are strong against the no probe control filter (Fig. 7(A), (B)). CY3-labeled probes were difficult to differentiate from autofluorescence but so strong that they bled into the FITC channel and high signals were detected. Positive results for *M. aeruginosa,* but not for *P. agardhii* were seen (Table 6). *Planktothrix* was recorded with the microarray. For *Microcystis* we can compare both results. In 2014, only three sampling times had a microarray result but no FISH result. In 2015, two times had a microarray result but no FISH result (Table 6). Microarray signals should be higher because 30 mls of concentrate were taken for RNA extraction. For FISH, two mls of raw water were filtered. For the highest microarray signals, there were fewer FISH labelled cells and vice versa (Fig. 7). A high microarray signal for prokaryotes indicates either many senescent cells or a few cells that are physiologically very active. FISH results would indicate the latter true for October 2014 and the former for 2015 samples. Up to 8% of the cells were *Microcystis* when many signals were recovered but the inferred cell counts would indicate that there were fewer cells present (Table 3), indicating presence only. Only single cells of *Microcystis* were found in the Tiber River.



(A) (B) (C)

Figure 7: (A) FISH hybridisation for *M. aeruginosa* seen through three different filters: DAPI, CY3 and FITC: a,b,c = no probe control, d,e,f = FITC labelled probe, g,h,i = CY3 labelled probe, j,k,l = both CY3 and FITC labelled probes. (B) FISH hybridisation for *P. agardhii* seen through three different filters: DAPI, CY3 and FITC: a,b,c = no probe control, d,e,f = FITC labelled probe, g,h,i = CY3 labelled probe, j,k,l = both CY3 and FITC labelled probes used together. (C) Representative images from the testing of the Tiber River samples with the species level probe for *Microcystis aeruginosa.* a. Positive control culture (note colony formation), 100X, b. Sample with isolated cells, c. Sample with many single cells, 10X.

Table 6: Comparison of FISH and microarray results.

Sample	FISH *Planktothrix*	Microarray *Planktothrix**	FISH *Microcystis*	Microarray *Microcystis*
T1 Oct 2013	–	nd	–	nd
T2 Oct 2013	–	nd	–	nd
T3 Oct 2013	–	nd	–	nd
T4 Oct 2013	–	nd	–	nd
T1 Mar 2014	–	–	–	–
T2 Mar 2014	–	–	–	+
T3 Mar 2014	–	–	–	+
T4 Mar 2014	–	–	+ (nq)	+
T1 Oct 2014	–	–	–	+
T2 Oct 2014	–	–	+ (nq)	+++
T3 Oct 2014	–	–	+ (nq)	+++
T4 Oct 2014	–	–	+ (nq)	+++
T1 Mar 2015	–	–	–	+
T2 Mar 2015	–	–	+ (8.4%)	+
T3 Mar 2015	–	+	+ (2.09%)	+
T4 Mar 2015	–	–	–**	+
T1 Oct 2015	–	–	+ (4.05 %)	+
T2 Oct 2015	–	–	–	–
T3 Oct 2015	–	+	+ (0.65%)	+
T4 Oct 2015	–	+	–	+
T3 January 2016	–	nd	+ (2.45%)	nd

*At the genus level, *Planktothrix* was common at most stations throughout the sampling period; nd: no data; nq: present but not quantified; ** impossible to differentiate fluorescence probe from fluorescence control, + fluorescence detectable, +++ normalised fluorescence signal > 3.

4 CONCLUSIONS

Our phylochip has been successfully applied to detect the presence of pathogens along the length of the Tiber River with FISH validation for cyanobacteria. Although our microarray has detected the presence of toxic cyanobacterial species, no toxin analysis was performed. In a study of cyanobacterial toxins associated with μAQUA samples, good correlation between toxins and species identified by the microarray was obtained [17], confirming toxins were retained by the hollow fibre filters. In μAQUA, we developed a reverse transcriptase probes extension for toxin genes direct on the microarray to detect toxin potential of any population (see [5]). We recommend using both phylochips for species and toxin detection systems for better human health protection because not all strains of a toxic species are equally toxic. Water authorities rely on turbidity and chlorophyll measurements as guides to possible toxic cyanobacterial events and our species microarray could help with mitigation.

REFERENCES

[1] Gentry, T.J., Wickham, G.S., Schadt, C.W., He, Z. & Zhou, J., Microarray applications in microbial ecology research, *Microbial Ecology*, **52**, pp. 159–175, 2006.

[2] Ye, R.W., Wang, T., Bedzyk, L. & Croker, K.M., Applications of DNA microarrays in microbial systems, *Journal of Microbiological Methods*, **47**, pp. 257–272, 2001.

[3] Kegel, J.U. Del Amo, Y. & Medlin, L.K., Introduction to project MIDTAL: its methods and samples from Arcachon Bay, France, *Environmental Science and Pollution Research*, 2013. DOI: 10.1007/s11356-012-1299-9.

[4] DeSantis, T.Z., Stone, C.E., Murray, S.R., Moberg, J.P. & Andersen, G.L., Rapid quantification and taxonomic classification of environmental DNA from both

prokaryotic and eukaryotic origins using a microarray, *FEMS Microbioloigcal Letter*, **245**, pp. 271–278, 2005.

[5] Kegel, J.U., Guillebault, D. & Medlin, L.K., Application of microarrays (phylochips) for analysis of community diversity by species identification. *Perspectives in Phycology*, in press, 2016.

[6] Wagner, M., Smidt, H., Loy, A. & Zhou, J., Unravelling microbial communities with DNA-microarrays, challenges and future directions. *Microbial Ecology*, **53**, pp. 498–506, 2007.

[7] Marcheggiani, S., D'Ugo, E., Puccinelli, C., Giuseppetti, R., D'Angelo, A.M., Gualerzi, C.O., Spurio, R., Medlin, L.K., Guillebault, D., Baudart-Lenfant, J., Weigel, W. Helmi, K. & Mancini, L. Detection of emerging and re-emerging pathogens in surface waters close to an urban area, *International Journal of Environmental Research and Public Health*, **12**, pp. 5505–5527, 2015.

[8] Jacob, P., Henry, A, Meheut, G., Charni-Ben-Tabassi, N., Ingrand, V. & Helmi, K., Health risk assessment related to waterborne pathogens from the river to the tap, *International Journal of Environmental Research and Public Health*, **12**, pp. 2967–2983, 2015. DOI: 10.3390/ijerph120302967.

[9] Panaiotov, S., et al., Two-year monitoring of water samples from Dam of Iskar and the Black Sea, Bulgaria, by molecular analysis: Focus on *Mycobacterium spp*., *International Journal of Environmental Research and Public Health*, **12**, pp. 7430–7443, 2015. DOI: 10.3390/ijerph120707430.

[10] Lewis, J., Medlin, L.K. & Raine, R., *MIDTAL (Microarrays for the Detection of Toxic Algae), A Protocol for a Successful Microarray Hybridisation and Analysis*, Koeltz: Koenigstein, Germany, 2012.

[11] Yilmaz, M., Philips, E., & Tillett, D., Improved methods for the isolation of cyanobacterial DNA from environmental samples, *Journal of Phycology*, **45**, pp. 517–521, 2009.

[12] Metfies, K., Borsutzki, P., Gescher, C., Medlin, L.K. & Frickenhaus, S., Phylochip Analyzer – a program for analysing hierarchical probe-sets, *Molecular Ecology Research*, **8**, pp. 99–102, 2008.

[13] Dittami, S. & Edvardsen, B., GPR-analyzer: a simple tool for quantitative analysis of hierarchical multispecies microarrays, *Environmental Science and Pollution Research*, 2012. DOI: 10.1007/s11356-012-1051-5.

[14] Caraux, G. & Pinloche, S., PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order, *Bioinformatics*, **21**, pp. 1280–1281, 2005.

[15] Töbe, K., Tullis, D., Gladstone, M. Anderson, D., & Medlin, L.K., Detecting intact algal cells with whole cell hybridisation assays. In: Karlson, B., Cusack, C. & Bresnan, E., (eds), *Microscopic and Molecular Methods for Quantitative Phytoplankton Analysis. IOC Manuals and Guides, No. 50*, Inter. Ocean. Com. UNESCO, pp. 55–66, 2010.

[16] Groben, R. & Medlin, L.K., *In situ* hybridisation of phytoplankton using fluorescently-labelled rRNA probes. In: Zimmer, E.A. & Roalson, E., (eds), *Methods in Enzymology*, Elsevier: San Diego, CA., pp. 299–310, 2005.

[17] Rodriguez, I., et al., Monitoring of freshwater toxins in European environmental waters by using novel multi-detection methods, *Environmental Toxicology and Chemistry*, in press, 2016.