Molecular ecological and phylogenetic analyses of the intestinal microbiota of the ascidian tunicates *Boltenia echinata, B. ovifera, Halocynthia pyriformis and Ciona intestinalis*

H. Rex Gaskins[1], Nathaniel R. McCray IV[2], Chad T. Collier[1], Dale E. King[1], Joel E. Thurmond[1], & Roderick I. Mackie[1]
[1]University of Illinois at Urbana-Champaign, Urbana, IL 61801
[2]Morehouse College, Atlanta, GA 30314

We are addressing the working hypothesis that enzymic detoxification (biotransformation) of xenobiotics by intestinal bacterial symbionts serves as a mechanism underlying their stable association with marine invertebrates. Here, we report results from initial studies that examined via 16S ribosomal (r) DNA-based approaches, the relative complexity and phylogenetic diversity of the intestinal microbiota of the North Atlantic ascidians (tunicates), *Boltenia echinata* (Cactus sea squirt), *B. ovifera* (Stalked sea squirt), *Halocynthia pyriformis* (Sea peach), *and Ciona intestinalis* (Sea vase) (Fig. 1). A question of special interest was the degree of host specificity in the composition of intestinal bacteria among these four tunicate species, which co-inhabit the sublittoral zone of the Gulf of Maine. The tunicates are of particular interest because they are: 1) sessile filter feeders, a lifestyle that might favor symbiosis with detoxifying microbes; 2) basal chordates, which diverged from other chordates before the large scale gene duplications that characterize vertebrate lineages; 3) *C. intestinalis* & *C. savignyi* genomes have been fully sequenced; and 5) they can be cultured *in vitro*, thus enabling toxicology and biochemical studies. A long-term goal is the identification of tunicate species that inhabit similar environments but differ in their carriage of detoxifying intestinal symbionts. Selected hosts can be used as comparative tools to test the hypothesis that development and support of a bacterial metagenome contributes to detoxification and has influenced, over evolutionary time, the host complement of biotransformation genes.

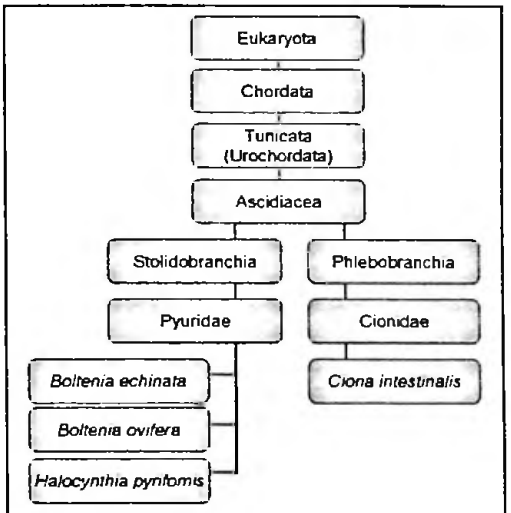

Fig. 1: Taxonomy of *B. echinata*, *B. ovifera*, *H. pyriformis* and *C. intestinalis*.

Until recently, microbial ecologists relied largely on techniques requiring cultivation of organisms on selective media. Although cultivation-based studies are useful for analysis of specific groups of bacteria, this approach is time- and labor-intensive, and the use of selective media imposes a presumptive bias on the types of bacteria that can be identified. Furthermore, only 10-20% of intestinal bacterial species may be represented using cultivation-based techniques[11]. These limitations have been now partly circumvented by the introduction of higher resolution molecular techniques to analyze microbial community diversity using 16S rRNA or rDNA as a molecular fingerprint to identify and classify organisms[15]. Polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE) is a technique in which DNA is isolated from a mixed sample and amplified using conserved 16S rDNA domain (Bacteria and Archaea) primers[6]. Although all PCR products are of approximately equal size, when electrophoresed on a polyacrylamide gel containing an increasing gradient of DNA denaturants, individual amplicons cease to migrate as the double-stranded products denature according to their G/C content. As a result, this approach allows separation of individual sequences based on G/C content as they correspond to the different bacterial species within the sample. Gel banding patterns from mixed samples can be compared to evaluate the relative similarity of bacterial communities from different habitats and host species. In addition, after electrophoresis, individual

PCR amplicons can be excised from the gel for sequencing and phylogenetic identification. Thus, 16S rDNA PCR-DGGE and combined sequence analyses of the intestinal microbiota enables an objective comparison that is unbiased by a presumptive decision on the types of bacteria to be analyzed

Replicate specimens of *B. echinata, B. ovifera, H. pyriformis, and C. intestinalis* were collected from Cobscook Bay by Gulf of Maine Marine Life Supply Company (http://www.gulfofme.com) staff and shipped overnight in sea water to MDIBL. The intact intestine was dissected within 24 hours of specimen collection according to dissection guides provided by the Online Invertebrate Lab from Lander University (http://www.lander.edu/rsfox/310labindex.html). DNA was isolated from combined mucosal scrapings and luminal contents using an UltraClean™ Soil DNA Kit (MO BIO Laboratories; Carlsbad, CA; http://www.mobio.com). 16S-V3 rDNA PCR-DGGE and community structure analyses, and the cloning and sequencing of individual amplicons were performed as described previously[5,6]. A BLAST search of the Entrez Nucleotides Database (http://www.ncbi.nlm.nih.gov/BLAST/) was performed to determine the taxa with the closest 16S V3 rDNA sequence match to the amplicons (bands) analyzed from PCR-DGGE. Positive identification of unknown sequences was considered significantly similar when 97 to 99% identical to BLAST database sequences.

A representative 16S V3 DNA PCR-DGGE gel image for replicate hosts is presented in Fig. 2*A*. Bacterial profiles were least diverse for the two *Boltenia* species with *B. echinata* (n = 3) harboring an average of eight 16S V3 rDNA amplicons (~ bacterial species) and *B. ovifera* (n = 3) 10 amplicons,

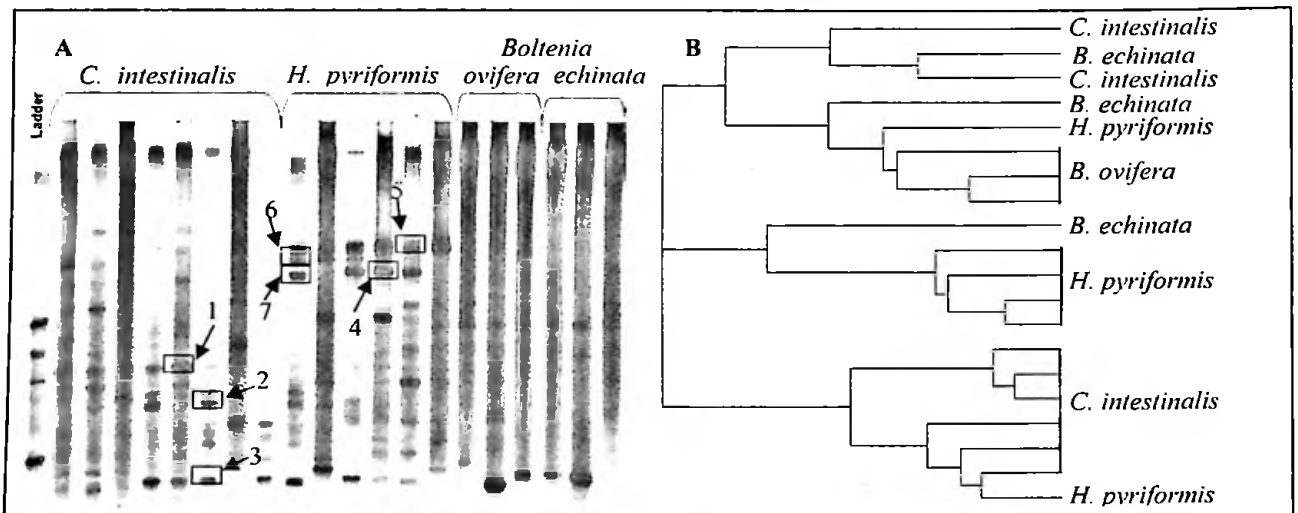

Fig 2. Molecular analysis of intestinal bacterial 16S-V3 rDNA isolated from *C. intestinalis, H. pyriformis, B. echinata*, and *B. ovifera*. A) PCR-DGGE profiles with a ladder of bacterial species with known migration distances. Sequenced bands are marked and identified in Table 1. B) Dendrogram of 16S-V3 rDNA banding patterns shown in (A). The dendrogram was constructed using Ward's algorithm[14] and Diversity Database software (Version 2.2.0; BioRad).

intermediate for *H. pyriformis* (n = 6) with an average of 17 amplicons, while *C. intestinalis* exhibited the greatest bacterial diversity with an average of 21 amplicons among the eight replicate animals. A dendrogram of 16S V3 DNA banding profiles within and among host species was constructed via cluster analysis using Ward's algorithm[14] (Fig. 2*B*). These data demonstrate that bacterial communities varied little within each tunicate species, but variability among species was substantial. To our knowledge, these data are the first to describe the community structure of the intestinal microbiota for any tunicate species, as well as the trend for host specificity in this taxon. The latter finding is consistent with increasing evidence from 16S rDNA-based studies for host specificity in animal-microbe associations among a variety of vertebrate and invertebrate species[1,10,15].

Individual 16S rDNA amplicons were excised from DGGE gels and cloned and sequenced to phylogenetically identify distinguishing and common intestinal bacteria of *H. pyriformis* and *C. intestinalis* (Table 1). These results demonstrate that *H. pyriformis* uniquely harbored bacteria most closely related to various members of the Firmicutes phylum, which are included in a restricted group of bacterial taxa that possess cytochrome P450 monooxygenases (P450s)[2]. The P450s constitute one of the most diverse enzyme superfamilies, and are involved in a variety of both anabolic and catabolic metabolic processes[4]. Although less well-characterized than the mammalian P450s, some bacterial P450s have been shown to be involved in the biosynthesis of antimicrobials and oxidative biotransformation of natural compounds or man-made chemicals[4].

Table 1. Summary of 16S rDNA sequences cloned from the intestinal microbiota of *C. intestinalis and H. pyriformis*

Band[A]	Source[B]	Closest Taxon[C]	Phylum, Order	% Similarity[D]
1	**Ci**	*Aurantimonas coralicida* (AY065627)	Proteobacteria (Alpha), Rhizobiales	172/196, (88%)
2	**Ci**	*Shewanella sp.* (AJ563805)	Proteobacteria (Gamma), Alteromonadales	194/196 (99%)
3	**Ci**, Hp	Uncultured α proteobacterium (AF509578)	Proteobacteria (Alpha), unclassified	191/196 (97%)
4	**Hp**	*Entomoplasma lucivorax* (AF547212)	Firmicutes (Gram⁻ bacteria), Entomoplasmatales	172/196 (88%)
5	**Hp**	Uncultured bacterium clone (AY171315)	Unclassified, from marine sediment	169/196 (86%)
6	**Hp**	*Leuconostoc pseudomesenteroides* (X95979)	Firmicutes, Lactobacillales	169/196 (86%)
7	**Hp**	*Mycoplasma collis* (AF538681)	Firmicutes, Mycoplasmatales	51/196 (26%)

[A]PCR-DGGE bands as numbered in Fig. 2.
[B]Ci, *Ciona intestinalis*; Hp, *Halocynthia pyriformis*. Bold indicates host of sequenced clone.
[C]Taxon and corresponding accession number of the closest 16S-V3 rDNA sequence match to the cloned amplicons (bands) as determined via a BLAST search of the Entrez Nucleotides database (http://www.ncbi.nlm.nih.gov/BLAST/).
[D]Percent similarity of sequenced 16S-V3 PCR-DGGE amplicons to sequence of closest taxon.

Each of the three amplicons sequenced from *C. intestinalis* were most similar to various proteobacteria. Relevant to our working hypothesis, it is intriguing that the picture now emerging from sequenced bacterial genomes is that many proteobacteria contain large sets of widely divergent genes encoding glutathione S-transferase (GST), a key enzyme in phase II biotransformation reactions in eukaryotes[13]. This suggests the possibility of a complex history of gene loss, replacement and acquisition by horizontal transfer of GST genes among the diverse proteobacteria. The few bacterial GSTs for which substrates are known, are catabolic enzymes with an essential role for growth on recalcitrant chemicals, rather than being conjugative enzymes, the typical function of eukaryotic GSTs[13]. For the most part, these were identified as GSTs only after glutathione was shown to be required for growth or the corresponding gene was sequenced[12]. Much additional work is required to better define the diversity, role and potential evolutionary influence of the bacterial P450s and GSTs. To phylogenetically identify *C. intestinalis* and *H. pyriformis* symbionts harboring P450 and GST genes, we will pursue a metagenomic approach that enables the cloning of large genomic fragments from mixed microbial community DNA[8]. The vector used in this approach is designed to contain, in each cloned fragment, both protein-encoding sequences and a phylogenetically-informative end, thus enabling direct linkage of specific functions with the responsible phylogenetic groups in the mixed microbial community.

On the host side, comparative analysis of the human (http://www.ncbi.nlm.nih.gov/genome/guide/human/) and *C. intestinalis* (http://genome.jgi-psf.org/ciona4/ciona4.home.html) genomes revealed 392 and 94 cytochrome P450 genes, respectively. Thus, the cytochrome P450 content of the human genome is approximately four-fold greater than that of *C. intestinalis*. Moreover, the GST (alpha, mu, and omega classes) content of the human genome (214 entries) is approximately seven-fold greater than that of *C. intestinalis* (32 entries). The 160 million base pair *C. intestinalis* genome contains approximately 16,000 protein-coding genes, which is similar to the number found in the model organisms *Drosophila melanogaster* (~14,000) and *Caenorhabditis elegans* (~19,000), but roughly half the number estimated to be present in the human genome (~30,000)[3]. Consistent with the reduced content of P450 and GST genes, most genes present as multiple copies in vertebrates only have a single representative in *Ciona*[9]. We are now pursuing phylogenomic approaches to gain insight into the potential role of the *Ciona* P450 and GST genes. It is predicted that this information together with knowledge of the P450 and GST gene families in their bacterial symbionts will verify that *Ciona* is an additional, easily accessible model organism that will enable a better understanding of the phylogenomics and functions of biochemical detoxification systems.

Supported by a New Investigator Award (HRG) with funds from the Salisbury Cove Research Fund and MDIBL's NIEHS Center for Membrane Toxicity Studies (ES03828-18). Nathaniel R. McCray IV was the recipient of a summer student fellowship from MDIBL NSF Research Experience for Undergraduates Site (DBI DBI-0139190). Su-Young Park (University of Illinois) is acknowledged for her assistance with 16S rDNA cloning as are staff of the W. M. Keck Center for Comparative and Functional Genomics (University of Illinois) for sequence analysis.

1. Ashen JB and Goff LJ. Molecular and ecological evidence for species specificity and coevolution in a group of marine algal-bacterial symbioses. *Appl Environ Microbiol* 66: 3124-3030, 2000.
2. Degtyarenko KN and Kulikova TA. Evolution of bioinorganic motifs in P450-containing systems. *Biochem Soc Trans* 29: 139-147, 2001.
3. Dehal P, Satou Y, Campbell RK, et al. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298: 2157-2167, 2002.
4. De Mot R and Parret AHA. A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. *Trends Microbiol* 10: 502-508, 2002.
5. McCracken VJ, Simpson JM, Mackie RI, and Gaskins HR. Molecular ecological analysis of dietary and antibiotic-induced alterations of the mouse intestinal microbiota. *J Nutr* 131: 1862-1870, 2001.
6. Muyzer G, Brinkhoff T, Nubel U, Santegoeds C, Schafer H and Wawer C. Denaturant gradient gel electrophoresis in microbial ecology. In: Molecular Microbial Ecology Manual, Vol. 3.4.4 (Akkermans A, van Elsas JD and de Bruijn F, eds), pp. 1-27. Kluwer Academic Publishers, Boston, MA, 1998.
7. Nebert DW and Gonzalez FJ. P450 genes: structure, evolution and regulation. *Annu Rev Biochem* 56: 945-993, 1987. Rondon MR, August PR, Bettermann AD, et al. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* 66: 2541-2547, 2000.
8. Satoh N, Satou Y, Davidson B, and Levine M. *Ciona intestinalis*: an emerging model for whole-genome analyses. *Trends Genet* 19: 376-381, 2003.
9. Taylor ME, Schupp PJ, Dahllöf I, Kjelleberg S, Steinberg PD. Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity. *Environ Microbiol* 6: 121-130, 2004.
10. Vaughan EE, Schut F, Heilig HGHJ, Zoetendal EG, de Vos WM and Akkermans ADL. A molecular view of the intestinal ecosystem. *Curr Issues Intest Microbiol* 1: 1-12, 2000.
11. Vuilleumier S. Bacterial glutathione *S*-transferases: what are they good for? *J Bacteriol* 179: 1431-1441, 1997.
12. Vuilleumier S and Pagni M. The elusive roles of bacterial glutathione *S*-transferases: new lessons from genomes. *Arch Microbiol Biotechnol* 58: 138-146, 2002.
13. Ward JH. Hierarchical grouping to optimize an objective function. *J Amer Statist Ass* 58: 236-244, 1963.
14. Zoetendal EG, Collier CT, Koike S, Mackie RI, and Gaskins HR. Molecular ecological analysis of the gastrointestinal microbiota: a review. *J Nutr* 134: 465-472, 2004.