

The Search for Novel Integron Expression Along a Heavy Metal Gradient

Brian Suh¹, Anne Giblin², Julie Huber³

Department of Biology¹, Haverford College, 370 Lancaster Avenue, PA 19041

Ecosystems Center², Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543

Josephine Bay Paul Center³, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543

Abstract

Since the discovery of antimicrobials, many life threatening illnesses have diminished in severity. This victory over bacterial infection, however, has shown itself to be only temporary. Bacteria have evolved to become much hardier, specifically due to the advent of multi-drug resistance. Using a platform known as integrons, bacteria are able to excise and insert gene cassettes encoding for resistance into their own genomes. These gene cassettes can be taken from their surrounding environment, often following conjugation with other bacteria. This horizontal gene acquisition process, distinctly different from vertical parent-offspring gene transfer, greatly increases the rate of dissemination of genetic material and rapidly expands the spectrum of antimicrobial resistance. As the emergence of resistant bacteria was discovered in clinical settings, many of the genes found in integrons to date have coded for resistance to drugs. This potent genetic platform has recently been found to have significant applications outside of the clinical setting in natural environments. I set out to analyze the DNA sequences of bacteria living in an environment of selective stress in the form of heavy metals. Doing so would provide insight into the diversity of genes found in environmental integron samples, as well as possibly novel gene products being used by bacteria to cope with stress inducing levels of heavy metals. As increased heavy metal concentrations should correlate with increased stress on microorganisms, increased integron expression of genes related to combating this stress should be observed. Desired integron DNA extracted from environmental sediment samples were isolated and amplified using PCR. These products were then ligated into plasmid vectors and electroporated into *Escherichia coli* for cloning. These genes were then extracted and sequenced to be analyzed by cross-referencing sequence data with the Basic Local Alignment Search Tool (BLAST). Unfortunately, the vast majority of this experiment failed to produce any significant results. A variety of issues dealing with DNA extraction and PCR severely hampered forward progress, and sequencing was done with a low degree of confidence. Final sequence data was unusable or inconclusive.

Key Words/Phrases: integron, environment, sediment, heavy metals, ligation, electroporation, cloning, Basic Local Alignment Search Tool (BLAST)

Introduction

In the last ten to fifteen years, treatment of bacterial infection using antibiotics has become more difficult. The widespread use of antibiotics to treat infections, especially in the United States, has resulted in the accelerated evolution of bacteria. Many bacteria have acquired resistance to more than one drug, a feat once thought to be beyond the grasp of bacterial evolution (Bennett, 2008). This new property can be attributed to horizontal, rather than vertical gene transfer, through the activity of integrons.

These mobile genetic elements are composed of distinct conserved regions that enable researchers to identify, isolate, and characterize integrons of both known and unknown origin and function (Bennet, 2008). Every integron begins with a highly conserved 5' region, containing the *IntI* gene, which acts as a promoter for incoming gene cassettes, as well as the *attI* gene, the promoter site for the transcription of these cassettes (Fig. 1). Furthermore, a common GTTRRRY sequence has been discovered within the *attI* site that is also represented in the 59-base elements (be) found downstream of each gene cassette. This sequence, particularly between the G and T, is where recombination crossover of cassettes takes place (Nield et al., 2001).

The integron is then capped by another conserved region at the 3' end. The region between the 5' and 3' conserved ends is highly variable, and can contain any number of integrated genes. Each plasmid that could potentially be inserted into an integron must contain the counterpart of the *attI* gene, *attC* (Mazel, 2006). These two genes interact to allow the new cassette to be incorporated, and allows us to distinguish between separate cassettes. This information will allow us to isolate integrons and their characteristic sub-units from DNA for further analysis and characterization.

Class 1 integrons specifically contain two resistance genes, *qacEΔ1* and *sul1*, which encode for resistance to quaternary ammonium compounds and for resistance to sulphonamides, respectively, towards this 3' conserved segment (Bennett, 2008). Class 1 integrons also tend to have nearly identical *intI* and *attI* genes (Nield et al., 2001). These were originally derived from transposon Tn402, and represent the vast majority of integrases found in clinical isolates (Mazel 2006). Class 2 and 3 integrons are specific to *Vibrio* bacterial strains, and are also heavily composed of antibiotic resistance genes. Class 4 integrons are found specifically on the small chromosome of *Vibrio cholerae*. The class 5 *intI* gene (*intI5*) has only been partially sequenced, but has been found to be associated with *Vibrio mimicus* strains and has been reported as being roughly 75% identity with *intI4* (Nield et al., 2001). These classes of integrons are mostly of clinical isolate or anthropogenic antimicrobial origin.

While useful and important, environmental integrons exist in an entirely different world, found in pristine environments or in those with unique selective stresses. Samples from these sources have yielded novel integrase genes that fall under new classifications (Nield et al., 2001). These environmental integrases are also noticeably different from clinically isolated integrases, showing less than 25% sequence homology. Thus, specific primers can be designed and utilized to focus attention on various classes of integrons. Further characterization of a variety of integrons and cassettes should allow for even more specific sample isolation and could also lend towards discovering new classes and gene cassettes. Some cassettes which tend to be too short for optimal PCR detection but were inadvertently discovered through their inclusion in larger integrons are an example of what advanced primer design could conceivably select for.

The field site chosen for this particular experiment is Great Sippewissett Marsh. This salt marsh is located near the Marine Biological Laboratory, on the Southern end of Cape Cod, Massachusetts. Experimental fertilization plots were established in April of 1974, onto which

sterilized sewage sludge was broadcast periodically (Valiela, 1975). The frequency of these fertilizations, as well as its composition, has varied over time. Various dosage levels were established, including a high dosage (HF) at $50.5 \text{ g}\cdot\text{m}^{-2}\cdot\text{wk}^{-1}$ and a low dosage (LF) $16.8 \text{ g}\cdot\text{m}^{-2}\cdot\text{wk}^{-1}$. The plots have a 10m radius (Fig. 2). Extra high (XF) fertilization plots spraying $151 \text{ g}\cdot\text{m}^{-2}\cdot\text{wk}^{-1}$ were established in 1974 (Harrold, 2005). Control plots are also present, in which no fertilizer has been added. Sediments taken from these four plots will allow me to establish a gradient of heavy metal concentration on which to compare integron sample characteristics, assuming that the environmental conditions affecting each of these plots and their heavy metal attenuation is equal.

Methods

DNA isolated from environmental sediment samples were processed, with gene cassettes cloned and sequenced individually. Preliminary PCR using select primer sets were used to indicate what classes of integrons are present, while other PCR performed further downstream were used to indicate the potential size of the gene cassette library, as well as to confirm the success of the final restriction enzyme analysis step before ultimately submitting samples for sequencing. These sequences were analyzed *in silico* using BLAST to look for general patterns and integron content.

Sediment samples were collected in triplicate from each experimental site at Great Sippewissett Marsh. These sites are designated as C, LF, HF, and XF to represent control, low fertilizer, high fertilizer, and extra high fertilizer, respectively. Each triplicate collection was done from sediment patches within one meter of each other, closer to the center of each experimental plot, to minimize variability. The triplicate samples were then combined into individual 50mL centrifuge tubes. Sediment samples were stored at -80°C . DNA was extracted from sediment samples using the Mo-Bio UltraClean® DNA Isolation Kit following the manufacturer's protocol. Extracted DNA was stored at -20°C .

For heavy metals analysis, I added approximately 0.1g of dry sediment from each site to individual 50mL centrifuge tubes. I then added 5mL of 12N concentrated HNO_3 to each of these tubes, as well as to two 'blank' tubes, for a total of 7 samples. These tubes were heated in an open water bath at roughly 70°C for about 2 hours. I swirled them gently halfway through until the mixture was homogenous. After this first incubation period, the tubes were allowed to cool back to room temperature for about an hour. I then added 5mL of trace metal grade concentrated HCl to each tube and placed them back in the water bath for 1 hour. These tubes were again swirled halfway through. The tubes were again allowed to cool to room temperature after this second incubation period.

I then filtered each sample into 100mL volumetric flasks using GF/F Whatmann filter paper number 1 (Fischer Scientific, Fairlawn, New Jersey). The tubes and filter paper were rinsed 3-4 times with DI water to ensure maximal yield. DI water was used to bring the flasks up to full volume. These samples were run using an atomic absorption spectrophotometer to analyze for lead, copper, and zinc levels.

At this point, PCR was performed using three distinct primer sets. The first was with specific primers for the *IntI1* gene characteristic of class 1 integrons (Leverstein-van Hall et al., 2003). Primer sequences are 5'-TCT CGG GTA ACA TCA AGG-3' (*IntI1-F*) and 5'-AGG AGA TCC GAA GAC CTC-3' (*IntI1-R*). Amplification was done in a total volume of 16 μL . This contains 15 μL of a master mix consisting of Platinum® PCR Supermix (Invitrogen) and 1 μL of each primer in a respective 100:1 ratio, as well as 1 μL of template DNA. Amplification using a thermal cycler proceeded as follows: 2 min at 94°C , followed by 35 cycles of 30 s at 94°C , 30 s at 55°C , and 30 s at 72°C . This followed by a terminal extension step for 2 min at 72°C (modified from Leverstein-van Hall et al., 2003). Another PCR using hep35 (5' TGC GGG TYA ARG ATB TKG ATTT 3') and hep36 (5' CARCATGCGTRTARAT 3') primers were used in conjunction with the *intI1* PCR to

differentiate between class 1,2, and 3 integrons. Amplification using a thermal cycler proceeded as follows: 2 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 55°C, and 45 s at 72°C. This was followed by a terminal extension step for 2 min at 72°C (White et al., 2000). The Lev5' (5'-GGC ATC CAA GCA GCA AG-3') and Lev3' (5'-AAG CAG ACT TGA CCT GA-3') primer set was used, which targets the conserved regions found on either end of any given gene cassette (Lévesque et al., 1995). PCR using the Lev or HS primers serves two purposes. It both amplifies the DNA regions of interest amongst the vast quantities of irrelevant DNA also present in the sample, as well as splices out and clone individual gene cassettes. An example of the difficulties I encountered while trying to visualize these gels is given in figure 4.

This amplified DNA was then ligated into vectors using the Topo® TA Cloning® Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. Electroporation was then used to transform these vectors into electrocompetent cells for further amplification. Plating onto Kanamycin plates allows for selective growth of successfully integrated vector colonies.

From here, the processed DNA was sent for sequencing. Receiving this information begins the next phase of computational comparison and analysis. Using the Basic Local Alignment Search Tool (BLAST), *in silico* analysis of these individual sequences is possible. BLAST would allow me to identify genes that have previously been identified and characterized by earlier research, and will also allow me to flag novel genes that have yet to be studied.

Results and Expected Results

Heavy metals analysis was successful. Each plot receiving higher amounts of fertilizer exhibits higher levels of lead, copper, and zinc (fig. 3). Sediment from Eel Pond was also

Unfortunately, I was unable to generate any significant data relating to integrons. For a variety of possible reasons, the vast majority of the PCR assays attempted failed to successfully visualize after gel electrophoresis, including unsuccessful DNA extraction and faulty or contaminated primers. To exacerbate the issue, control DNA also tended to function only sporadically. Limited success was thought to be observed with DNA from the HF plot, as seen in figure 5. Sequence results show that this was false, as the sequences were either too short to be of any significance (under 50 base pairs), or encoded for irrelevant *E.coli* genomic DNA which was most likely a relic of the *E.coli* medium used for cloning.

The nature of integron construction, as well as the lack of a comprehensive, universal gene library make it difficult to predict exactly what types of genes would be present in my samples. Indeed, the cassettes that I had hoped to find and sequence are what would have been contributing to this comprehensive gene library. The rapid horizontal dissemination of genes using this platform implies that countless numbers and types of genes will be circulating among bacteria, but only those that can successfully take up integrons containing useful genes to combat stressors would survive. Thus, the majority of cassettes found should encode for gene products that are directly related to processes involved in adapting to or neutralizing the effects of selective stresses. However, literature reviews demonstrate that integrons contain multiple gene cassettes, not all of which may be useful in dealing with stress. Because integrons are typically transcribed in their entirety, these useless or nonsense cassettes would 'piggyback' the system and thus also be prevalent in samples.

It must also be noted that the limitations and bias that come with primer design and PCR protocol will naturally prevent the capture of every individual gene cassette present in a sample (Stokes et al., 2001). PCR tends to favor the amplification of shorter length amplicons, so smaller gene cassettes will appear to be more abundant. Primer design represents the selection of very specific regions of DNA, and as such inherently excludes undesired segments, possibly inadvertently.

The types of genes most prevalent should be indicative of the environments the samples are extruded from (Nemergut et al., 2004). For example, samples taken from sediments high in heavy metal concentration might contain an abundance of metallothioneins, a family of proteins that specializes in sequestration of heavy metal ions to prevent them from interfering with regular cell function. These proteins containing large amounts of Cysteine residues. The thiol groups associated with Cysteine act as excellent sites for the interception of heavy metals, preventing them from damaging other components of the cell.

Conclusions

The study of integron activity and expression is relatively young. Each new integrase discovered, and each new primer set or technique developed helps us to learn more about how integrons operate, but also reminds us of how little we know about this powerful gene transfer system. A shotgun approach to the study of integrons might be useful in generating a massive gene library, but it would be more practical to approach the study from a systems perspective, comparing smaller gene libraries from different settings against each other to glean functional and perhaps more applicable information. By studying integrons in a select few environmental settings, I hope to be able to supplement the broader knowledge base about integrons in the scientific community. An understanding of how integrons function could have far reaching effects down the line, from clinical antibiotic design to environmental bioremediation using engineered microbes.

A more recent study describes a new primer set with a very high success rate and increased specificity for environmental *IntI* genes. Using the HS298 (5'-TGGATCCCACRTGNGTRTADATCATNGT-3') and HS286 (5'-GGGATCCTCSGCTKGARCGAMTTGTTAGVC-3') primers, three new classes of classes of integrons were discovered (Nield et al, 2001). This primer set could perhaps be used in conjunction with or in lieu of the Hep primers to more specifically address the environmental integrons I hope to study.

Another more recent primer set may present an updated and accurate approach to isolating gene cassettes. The HS286 (5' GGG ATC CTC SGC TKG ARC GAM TTG TTA GVC 3') and HS287 (5' GGG ATC CGC SGC TKA NCT CVR RCG TTA GSC 3') primers target separate halves of the 59 base element which is found at the end of each separate cassette (Stokes et al., 2001). Stokes estimates that sequences derived from the use of these primers will only account for about 25% of the entire gene library present in sediment samples. This perhaps suggests that multiple rounds of cassette isolation and cloning should be performed to capture as many of the cassettes present in these samples as possible.

Despite the failure of this experiment, I believe that the overall design and methods are very much intact. Heavy metals analysis data shows that the gradient I wished to study is indeed present. The possible sources of error or contamination should be accounted for when this experiment is carried out again. The use of the new primer sets should also help to increase efficiency and yield. Furthermore, the use of a different DNA isolation kit, PowerSoil® rather than UltraClean® could help to increase the initial DNA yield (Mo Bio, Carlsbad CA). While very similar in principle, the manufacturer claims that the PowerSoil® kit could be more ideal for the sediment samples I am examining. While this trial was ultimately inconclusive, repeating the experiment with the proposed changes should yield valuable results.

Fig 1. Structure of an integron (adapted from Bennett, 2008). This figure shows all of the conserved regions of a typical integron. Not depicted are the variable gene cassettes that could be incorporated into the integron, which would be found after the *int* gene and before the 3' conserved segment.

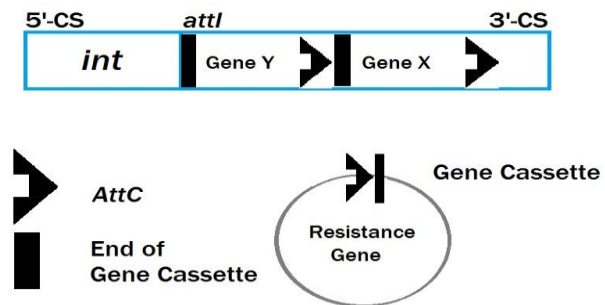


Fig 2. From Valiela et al. 1975. Site map of Great Sippewissett Marsh including experimental plots. LF and HF plots are shown, as well as control (C). Water level shown represents average tidal height. White spaces are covered in salt marsh vegetation, while Upland designates areas above the level of spring flood tides. Not depicted are the XF plots added after the original publication of this figure.

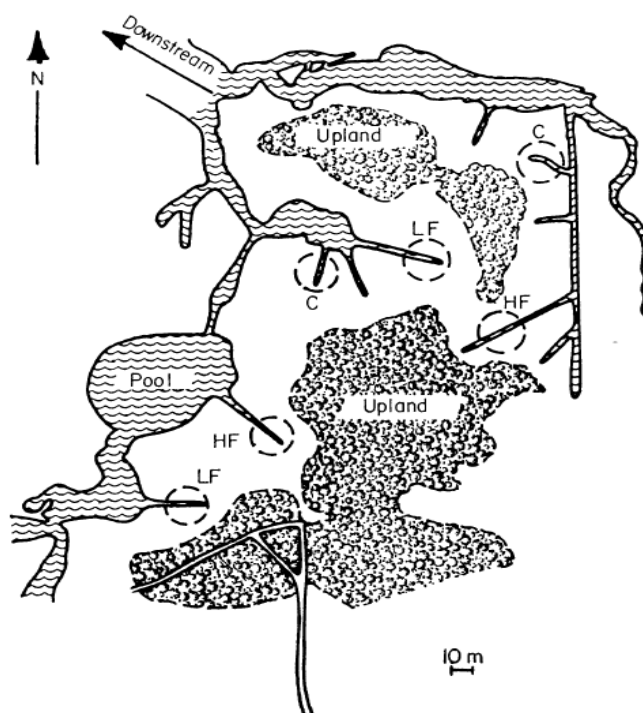


Fig 3. Heavy metal concentrations across each field site. As predicted, a gradient does exist at Great Sippewissett Marsh, which correlates to the amount of fertilizer being added at each experimental plot.

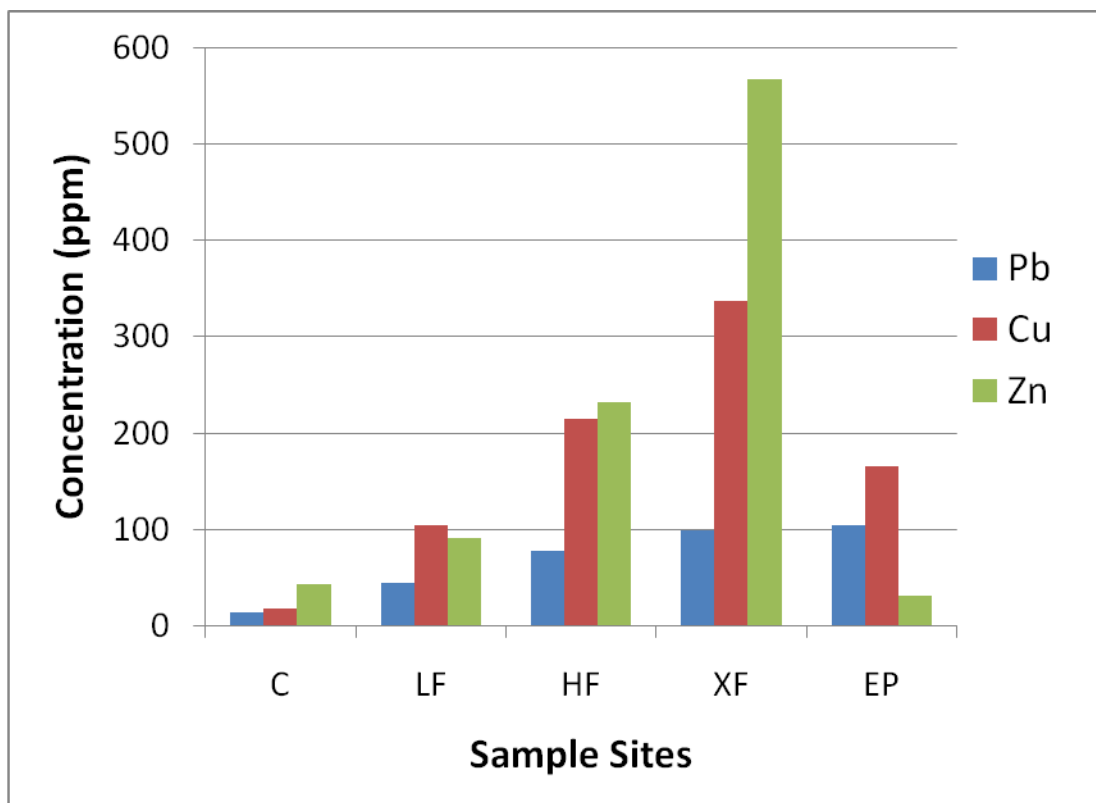


Fig 4. Gel electrophoresis showing an attempt to visualize PCR products in the search for the conserved *int* genes. Positive control DNA failed to amplify correctly, and experimental DNA also shows no signal.



Fig 5. Gel electrophoresis showing the results of *E.coli* cloning of individual gene cassettes. With a vector of known size (roughly 200 base-pairs), bands that visualize above this baseline indicate possible candidates for successful uptake of cassettes. These 15 samples were all individual colonies cloned from DNA extracted solely from the HF sediments.



Acknowledgements

I would like to thank the Julie Huber lab for providing me a workspace and supplies to carry out my experiment, especially Holly and Nuria for helping me become acquainted with the lab and providing a friendly and supportive atmosphere. Thank you so much to Anne Giblin for her guidance and giving me the idea of a heavy metal gradient on which to focus. I would like to thank Will Daniels for his aid in the heavy metals extraction. Finally, thanks to Julie Huber, for her unwavering confidence in my abilities, steadfast support, and encouraging personality.

Literature Cited

Bennett, P.M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British Journal of Pharmacology*. 153:S347-S357.

Leverstein-van Hall, M.A., Blok, H.E.M., Donders, A.R.T., Paauw, A., Fluit, A.C., and Verhoef, J. (2003). Multidrug Resistance among Enterobacteriaceae Is Strongly Associated with the Presence of Integrons and Is Independent of Species or Isolate Origin. *Journal of Infectious Diseases*. 187:251-9.

Lévesque C., Piché, L., Larose, C., and Roy, P.H. (1995). PCR Mapping of Integrons Reveals Several Novel Combinations of Resistance Genes. *Antimicrobial Agents and Chemotherapy*. 39:185-191.

Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nature*. 4:608-620.

Nemergut, D.R., Martin, A.P., and Schmidt, S.K. (2004). Integron Diversity in Heavy-Metal-Contaminated Mine Tailings and Inferences about Integron Evolution. *Applied Environmental Microbiology*. 70.2:1160-8.

Nield, B.S., Holmes, A.J., Gillings, M.R., Recchia, G.D., Mabbutt, B.C., Helena Nevalainen, K.M., and Stokes, H.W. (2001). Recovery of new integron classes from environmental DNA. *FEMS Microbiology Letters*. 195:59-65.

Stokes, H.W., Holmes, A.J., Nield, B.S., Holley, M.P., Helena Nevalainen, K.M., Mabbutt, B.C., and Gillings, M.R. (2001). Gene Cassette PCR: Sequence-Independent Recovery of Entire Genes from Environmental DNA. *Applied Environmental Microbiology*. 67:5240-6.

Valiela, I., J.M. Teal, and W. Sass. 1975. Production and dynamics of salt marsh vegetation and the effects of experimental treatment with sewage sludge: biomass, production and species composition. *Journal of Applied Ecology*. 12:973-982.

White, P.A., McIver, C.J., Deng, Y., and Rawlinson, W.D. (2000). Characterisation of two new gene cassettes, aadA5 and dfrA17. *FEMS Microbiology Letters*. 182:265-9.

White, P.A., McIver, C.J., and Rawlinson, W.D. (2001). Integrons and Gene Cassettes in the *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*. 45.9:2658-2661.