TRANSPOSABLE ELEMENT-LIKE REPEATS IN ALLERGY CAUSING GENES ISOLATED FROM SPF *Litopaneaus vannamei* POSTLARVAE EXPOSED TO 1 PPM OF CADMIUM – DEVELOPMENT OF ALLERGEN EST-SSRs FOR LINKAGE MAPPING

Alexandra Bowen[2]; Jacqueline Presedo[2], Laura Hake[2], Acacia Alcivar-Warren[1-2]

[1]Biology Department, Boston College, Chestnut Hill, MA 02467 USA,
[2]Environmental Genomics, Inc., Southborough MA 01772,
[3]FUCOBI (Fundación para la Conservación de Biodiversidad Acuática y Terrestre de Ecuador), Guayaquil, Ecuador

E-mail: environmentalgenomics.warren@gmail.com

ENVIRONMENTAL GENOMICS, INC.
IT'S ONE HEALTH

Litopenaeus vannamei

- Shrimp is the favorite seafood of Americans.
- Most shrimp is imported, creating a yearly $4.5 billion US trade deficit.
- *L. vannamei* is the most widely cultured species around the world, and the industry is impacted by viral diseases.
- *L. vannamei* bioaccumulates cadmium (Keating et al. 2007)
- A linkage map (*ShrimpMap2*) is being constructed to identify genes or Quantitative Trait Loci (QTL) associated with tolerance/resistance of shrimp to metals and persistent organic pollutants (POPs).

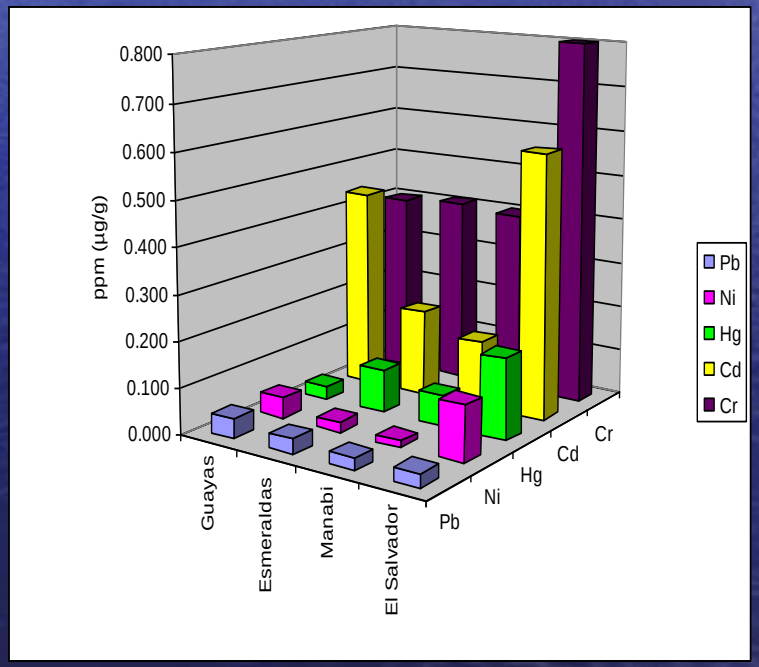
Shrimp allergies

- The number of people with shrimp allergies in the US has increased in the past 20 years, and most shrimp consumed by Americans is imported.

- Availability of specific pathogen-free (SPF) *L. vannamei*, the most widely farmed shrimp species around the world, has helped to control viral diseases and promote aquaculture both in the US and abroad.

Cadmium in *L. vannamei*

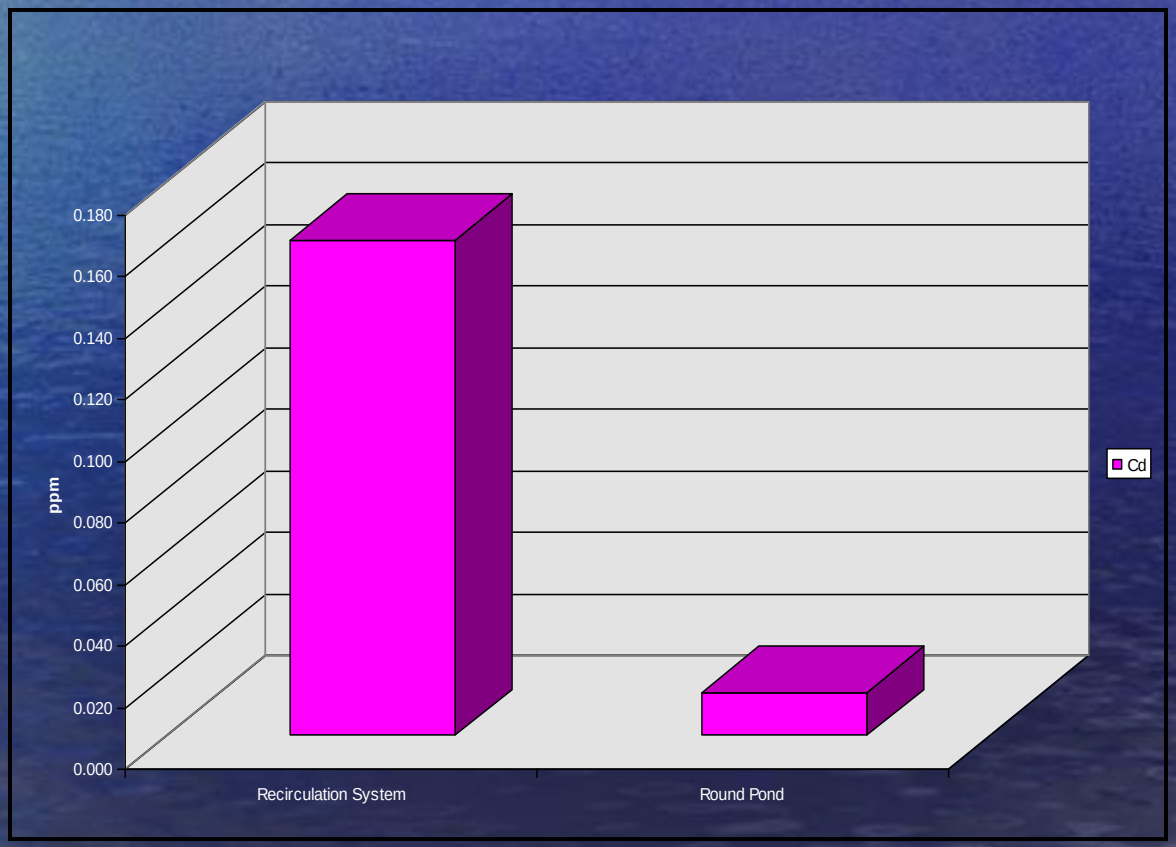
- 18 metals detected in wild shrimp from Latin America.
- Cadmium (Cd) is the metal of concern.
- Cd, present in crude oil, has negative health effects in most living organisms
- Cd accumulates in testes, prostate, renal epithelial cells, & is linked to:
  - kidney damage
  - prostate cancer
  - obesity and diabetes
  - learning disabilities
  - may affect immune & allergic responses



(Henry & Alcivar-Warren, 2004)

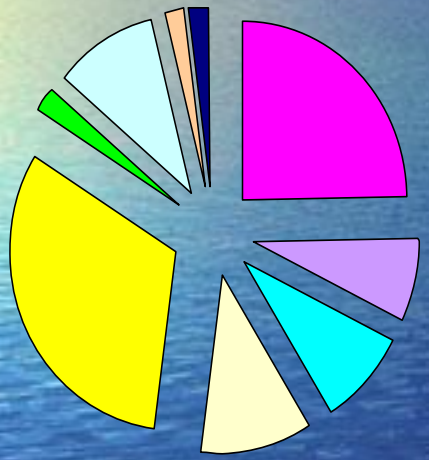
SPF shrimp bioaccumulate cadmium

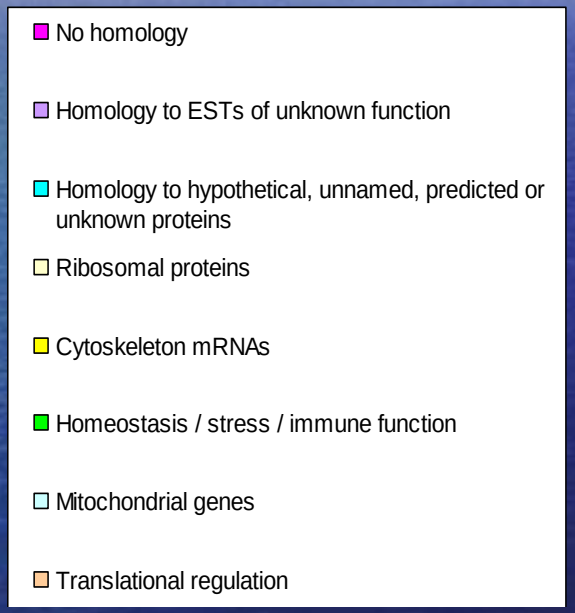
Cd levels increased 16-fold in Kona Line broodstock maintained in near zero-exchange recirculation system compared to a flow-through system.



(Alcivar-Warren & Meehan, 2004)

~3,000 allergy-causing genes isolated from TSV-challenged SPF *L. vannamei* juveniles

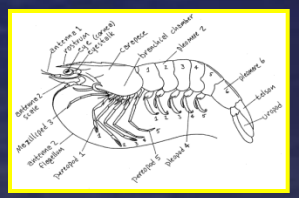




- Actin, myosin heavy chain, troponin I-C, tubulin,
- Tropomyosin, AK, MLC, SCP*
- Translationally controlled tumor protein (TCTP, or fortilin)
- Laminin receptor
- Chitinase, zinc proteinase
- heat shock protein 70
- antiviral protein hemocyanin,
- c-type lectins (PmAV & Kupffer cell receptor-like)
- 11.5 kDa antimicrobial peptide
- guanine nucleotide binding protein (rho pathway)
- nucleoside diphosphate kinase
- endoplasmic reticulum Ca2++
- 16s rRNA, COI, COII
- 60S, 40S ribosomal proteins, 28s, 18S, translation elongation & initiation factors
- ATPase, preamylase

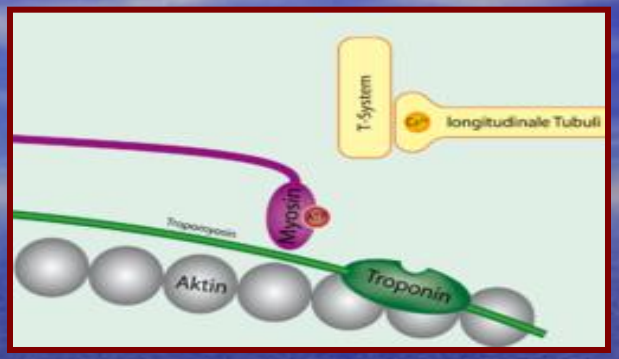
*known shrimp allergens

From Alcivar et al. 2011.

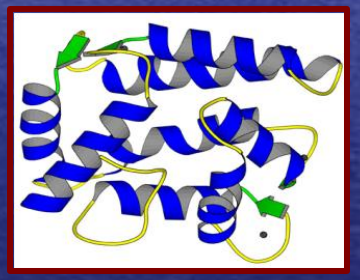


Shrimp Allergens

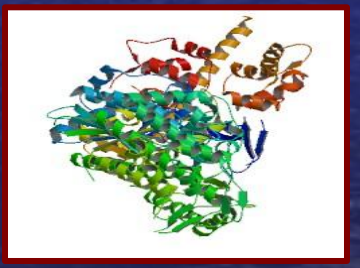
- Tropomyosin (TM)
  - -is involved in muscle contraction.
  - -along with **troponins,** it associates with the actin filaments and regulates actin mechanics
  - -is an **actin-binding** protein
  - -is regulated by the influx of **calcium ions**



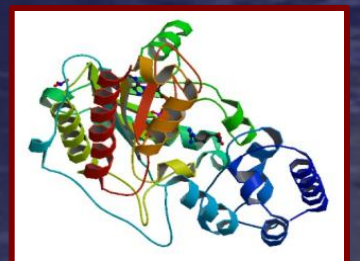
- Arginine kinase (AK)
  - -is a member of the phosphagen kinase family
  - -catalyzes the reversible transfer of a high-energy phosphate from the phosphagen arginine phosphate to ADP to form ATP.



- Myosin light chain (MLC)
  - ▪Smaller subunits of myosin that bind near the head groups of myosin heavy chains
  - ▪Involved in muscle contraction



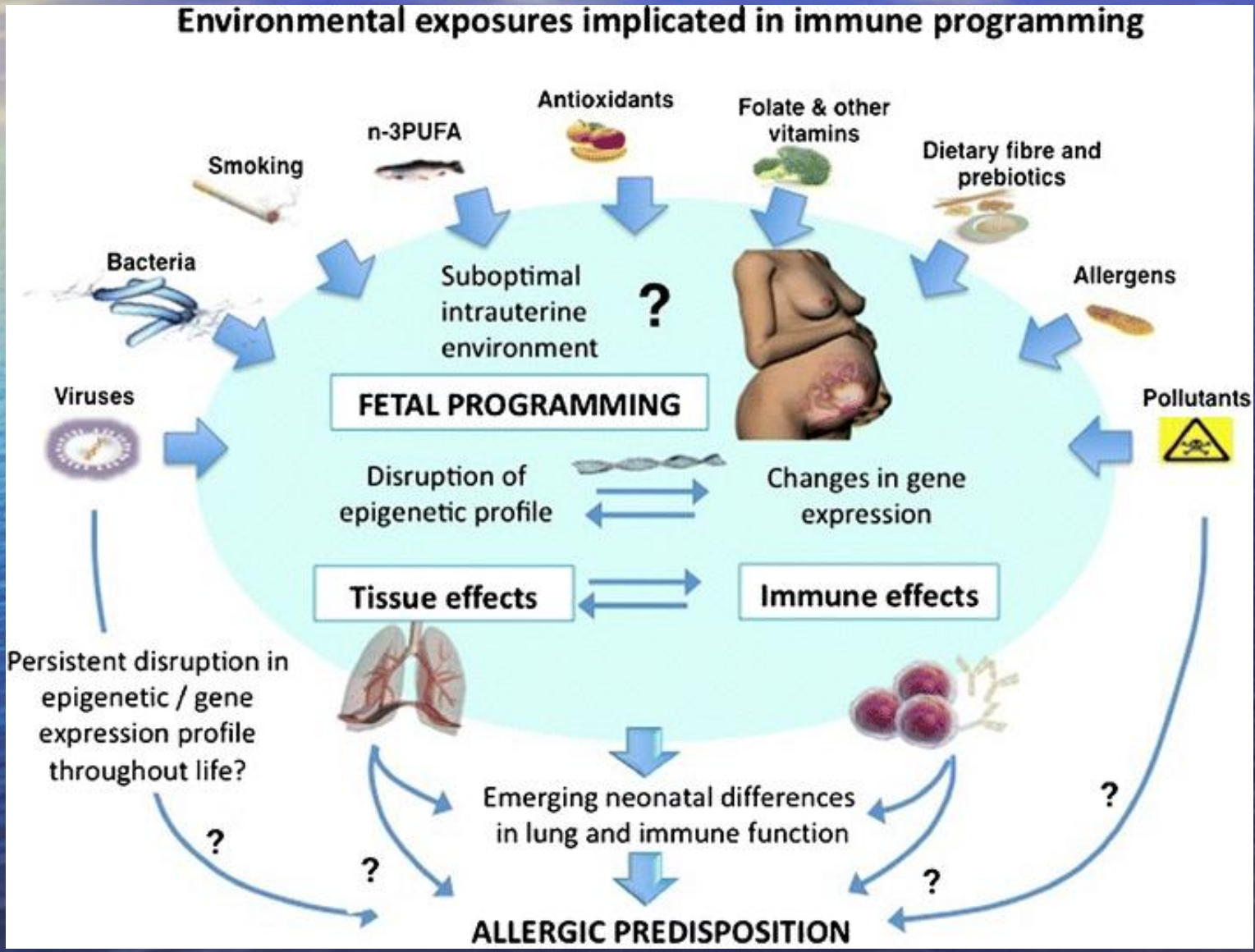
- Sarcoplasmic calcium-binding protein (SCP)
  - - is believed to function as the invertebrate equivalent of vertebrate parvalbumin, namely to "buffer" cytosolic Ca2+



There are ~4,000 allergen-related ESTs in the UniGene database – most are from SPF shrimp

Tropomyosin, TM Lit v 1	Arginine kinase, AK Lit v 2	Myosin Light Chain, MLC Lit v 3	Sarcoplasmic calcium-binding protein, SCP Lit v 4
Tropomyosin Lva.854: 80 seqs	Arginine kinase Lva.544: 984 seqs	Transcribed locus, moderately similar to NP_511049 myosin light chain cytoplasmic [D. melanogaster] Lva.1756: 183 seqs	Sarcoplasmic calcium-binding protein Lva. 2178: 644 seqs
		Transcribed locus, weakly similar to NP_001091813.1 myosin regulatory light chain 2 [Bombyx mori] Lva.1302: 906	Transcribed locus, weakly similar to XP_967547.1 similar to AGAP007963-PA [T.castaneum] Lva.20603: 21
		Transcribed locus, moderately similar to XP_976209.1 similar to myosin 1 light chain isoform 2 [Tribolium castaneum]. Lva.1350: 960	
		Lit v 3 allergen myosin light chain Lva.12166: 3 seqs **[2 are chimeric, with portions of both MLC2 & hypothetical protein]	
		Transcribed locus, weakly similar to NP_524586.1 myosin light chain 2, isoform A [D. melanogaster] Lva.6500: 20	
		Transcribed locus, weakly similar to XP_393544.3 PREDICTED: myosin light chain alkali-like isoform 4 [Apis mellifera] Lva.22925: 3	
		Transcribed locus, weakly similar to XP_001848910.1 myosin light chain kinase [Culex quinquefasciatus] Lva.12264: 8	

Seafood consumption as a factor in allergic predisposition?

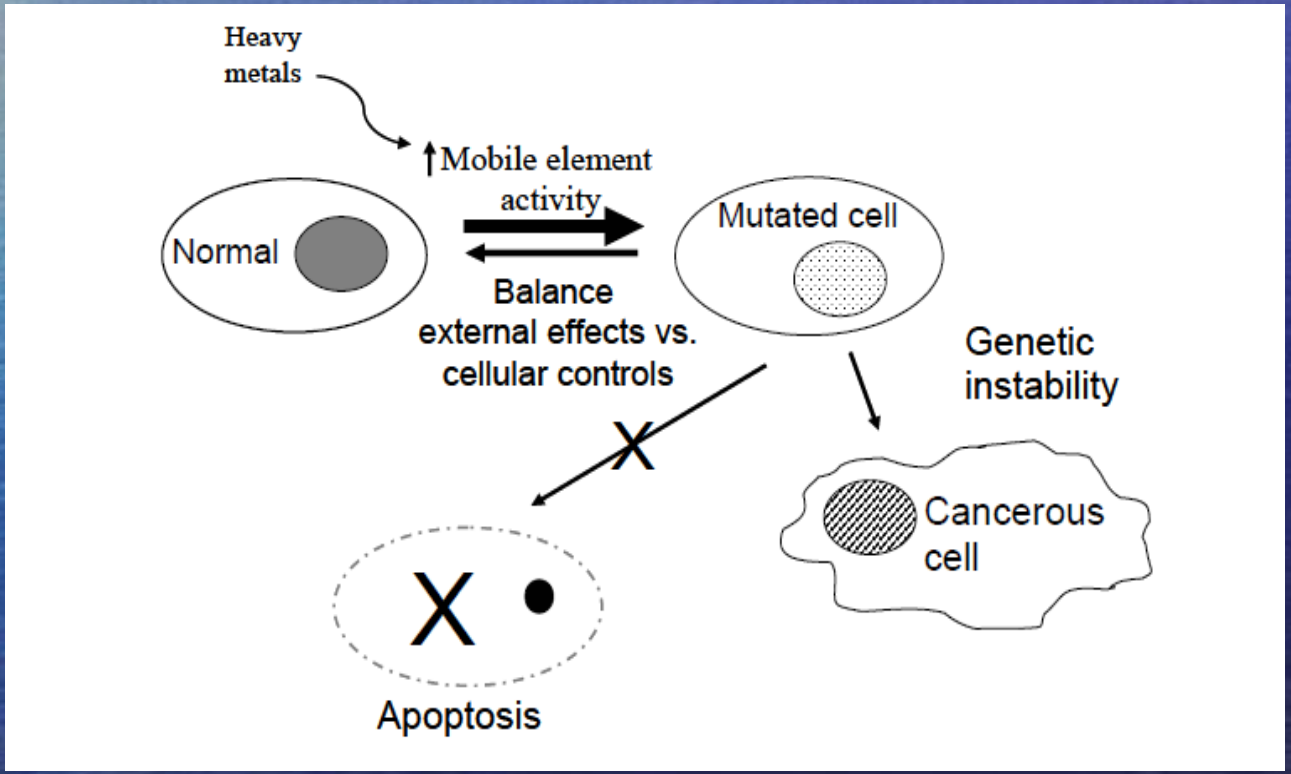


Transposable Elements in shrimp

- Differential expression of TEs in response to viruses and other environmental factors has been reported in *L. vannamei*, *L. stylirostris* (Hizer, 2007) and *P. monodon* (de la Vega, 2006; Tang & Lightner, 2006).
- The reverse transcriptase-like (RT-like) non-LTR retrotransposon has been identified in *P. monodon* carrying non-infectious sequences of IHHNV inserted in the genome (Tang & Lightner, 2006), or IHHNV-infected *L. stylirostris* (Hizer, 2007).
- Remnants of TE repeats have been identified in SPF *L. vannamei* challenged with TSV and WWSV (Alcivar-Warren, 2009; Das, 2009) or exposed to Cd (Keating, 2007).
- **It is possible that TEs increase the expression of allergens which are then are perceived as "foreign antigens" by hypersensitive people**, a hypothesis that merits testing.

Why study Transposable Elements?

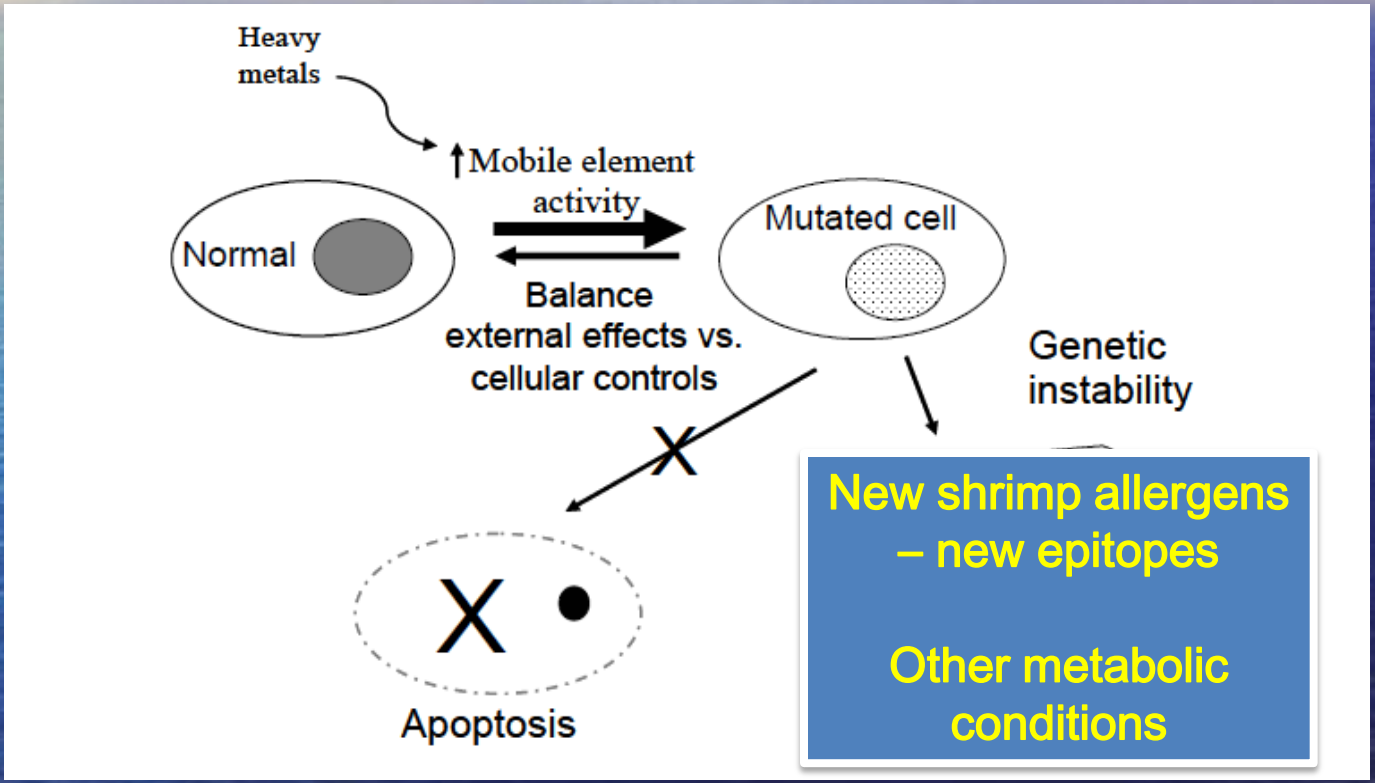
Because pollutants increase retrotransposon activity, which in turn affect fitness and disease susceptibility



LINE-1 non-LTR retrotransposon activity is induced by cadmium, mercury and nickel at very low levels, ppb (Kane et al. 2005)

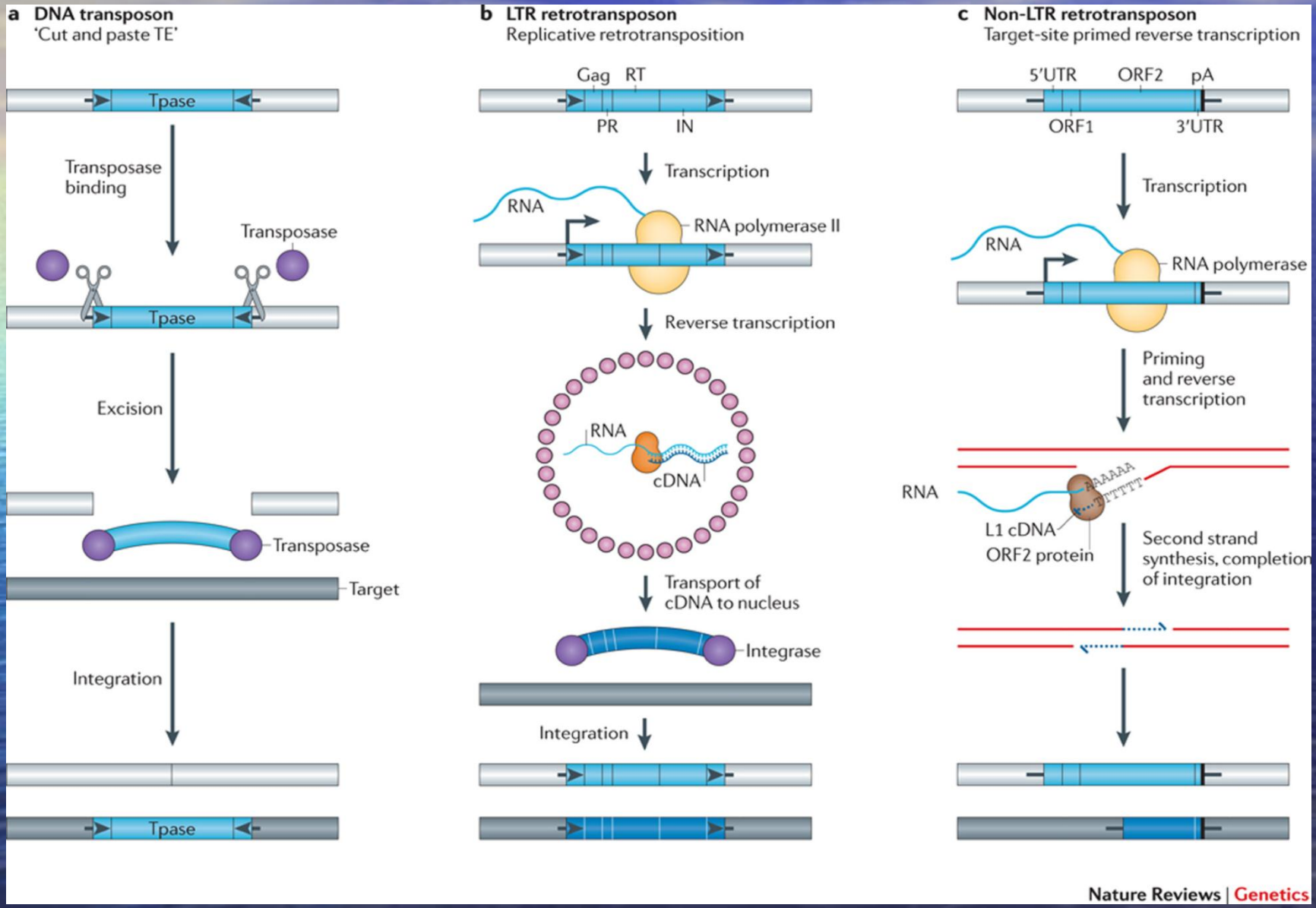
Hypothesis

Environmental stressors and viruses cause induction of transposable element activity, which cause host mutations



Is RTE-like, or LINE-1, or other non-LTR retrotransposons involved in virus infectivity, disease susceptibility?

Transposable element (TE)-like repeats are present in the shrimp genome



TE-repeats in myosin light chain-related ESTs

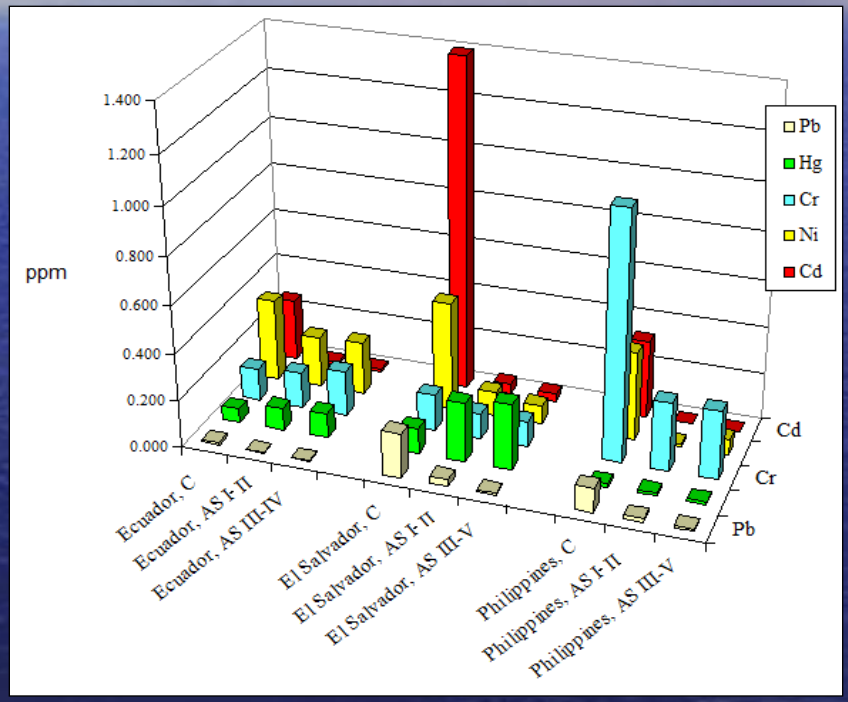
MLC - Lit v 3	Classification statistics using TEclass software	
	All ESTs	Control (Cd-PL42)
Transcribed locus, moderately similar to NP_511049 myosin light chainn cytoplasmic [D. melanogaster] Lva.1756: 183 seqs	DNA transposons: 1 LTRs: 52 LINEs: 65 SINEs: 1 Unclear: 64	
Transcribed locus, weakly similar to NP_001091813.1 myosin regulatory light chain 2 [Bombyx mori] Lva.1302: 906	DNA transposons: 35 LTRs: 508 LINEs: 166 SINEs: 4 Unclear: 193	0 (1) 1 (3) 0 (4) 0 (0) 6 (0)
Transcribed locus, moderately similar to XP_976209.1 similar to myosin 1 light chain isoform 2 [Tribolium castaneum] . Lva.1350: 960	DNA transposons: 62 LTRs: 168 LINEs: 324 SINEs: 19 Unclear: 387	0 (0) 1 (0) 1 (3) 0 (0) 3 (1)
Lit v 3 allergen myosin light chain Lva.12166: 3 seqs **[2 are chimeric, with portions of MLC2 & hypothetical protein]	LINEs: 3	
Transcribed locus, weakly similar to NP_524586.1 myosin light chain 2, isoform A [D. melanogaster] Lva.6500: 20	DNA transposons: 4 LTRs: 10 LINEs: 1 Unclear: 5	
Transcribed locus, weakly similar to XP_393544.3 PREDICTED: myosin light chain alkali-like isoform 4 [Apis mellifera] Lva.22925: 3	LTRs: 1 LINEs: 1 Unclear: 1	
Transcribed locus, weakly similar to XP_001848910.1 myosin light chain kinase [Culex quinquefasciatus] Lva.12264: 8	DNA transposons: 1 LTRs: 1 Unclear: 6	
Transcribed locus, weakly similar to XP_393371.3 PREDICTED: myosin regulatory light chain 2 [Apis mellifera] Lva.23243: 3	LTRs: 2 LINEs: 1	

Cadmium (Cd)

- No information on the potential risk of SPF shrimp on increased allergen expression has been reported.
- The toxic metal cadmium (Cd) is an important environmental pollutant, a carcinogen, a metallohormone, and an endocrine disruptor.
- Cd accumulates in testes, prostate, renal epithelial cells and liver, binds to DNA, and may cause heritable mutations.
- Chronic, sublethal Cd levels may be harmful to human health.
- Cd increases oxidative stress
- Cd induces the expression of transposable elements (TEs) such as LINE non-LTR retrotransposons
- may play a role in epidemics of diabetes worldwide, and exacerbates allergic reactions.
- Cd accumulates in the body of *L. vannamei* in proportion to environmental concentrations.

Cadmium (Cd), cont.

- Cd levels are 2-3X higher in cephalothorax (C) than abdominal segments (AS) of *L. vannamei* from Ecuador and Honduras and *Penaeus monodon* of Philippines, and may impact health of people that consume them.



Cadmium (Cd)

- Baseline expression data on the potential effects of Cd on shrimp health and growth suggested a complex interaction between environmental conditions and host genetic background.

- To begin to understand the molecular mechanisms involved in the increased prevalence of shrimp allergies, basic knowledge of shrimp genome structure is needed.

Hypothesis

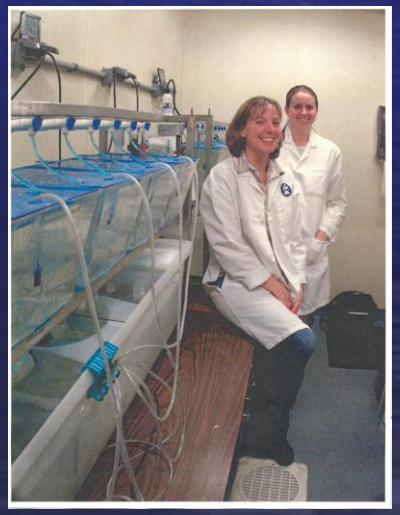
hypothesize that Cd induces the expression of allergy-causing genes and genome rearrangements mediated by TEs.

Objectives

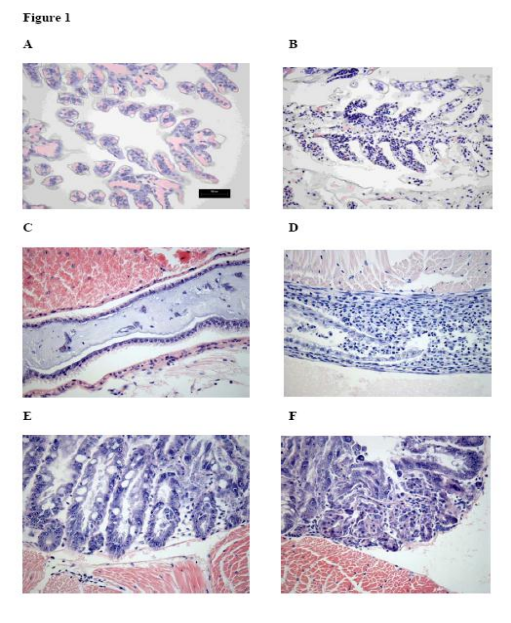
1. To determine if the number of allergen genes is increased after Cd exposure
2. To characterize the TE-like repeats in shrimp allergens, and
3. To develop EST-SSRs for mapping of shrimp allergens.

Experimental approach

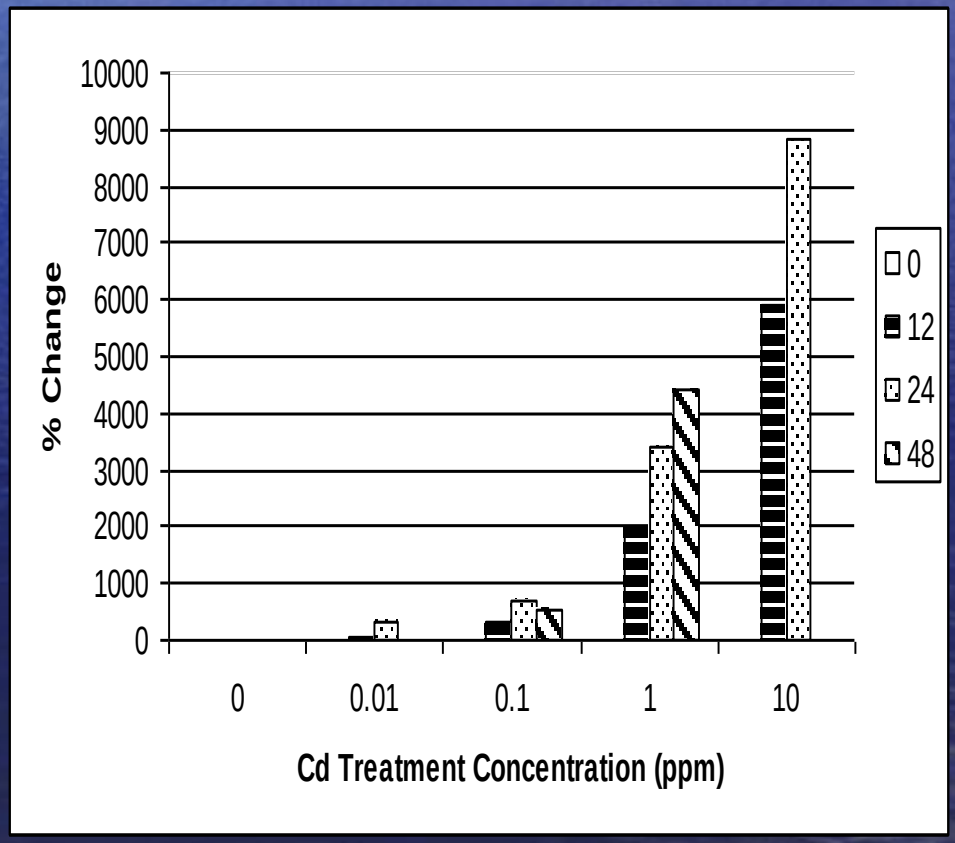
- Specific Pathogen Free (SPF) *L. vannamei* postlarvae (PL23) were obtained from the Kona Line of Oceanic Institute (OI), Kona, HI
- 11 day acclimation in Aquaneering (Ahab) system
- Artificial seawater
  - (salinity 23-26 ppt, pH 8.0)
- 12:12 light cycle
- 900 PLs used for assay (60 PLs per treatment, 3 replicates per treatment)



Cadmium bioaccumulation & histological changes in SPF *L. vannamei* postlarvae (PL42)



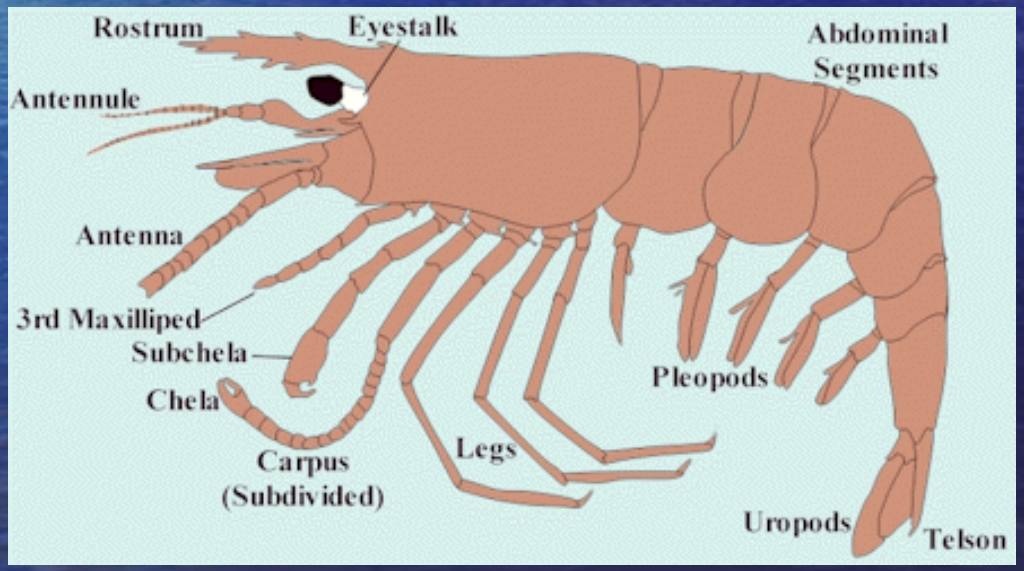
Histological changes in tissues of control (left) and Cd-treated PL42 (right).



Percent change of Cd levels in control and Cd-exposed SPF *L. vannamei* PL42

Methods for Objective 1:

- ESTs were isolated from cDNA libraries constructed using the cephalothorax of SPF, untreated *L. vannamei* postlarvae stage 42 (PL42) and Cd-treated (1ppm) postlarvae (Cd-PL42).



Objective 2 – Characterize the TE-like repeats in ALLERGEN genes

- Recombinant clones were sent to Amplicon Express, WA for cDNA library construction
- Agencourt, Inc. sequenced approx. 1500 clones from each library
- EST sequences were used for homology searches against Genbank database
- Bioinformatics: CENSOR and TEclass software
  - http://www.girinst.org/censor/index.php

RESULTS - Objective 1: Allergen genes in control and Cd-treated PL42

- There were no differences in the number of allergen genes isolated from control and Cd-treated cDNA libraries from cephalothorax

Objective 1. Gene Expression in Control and Cd-treated PL42

1. Differential gene expression analysis showed that 60.7% (622/1025) of ESTs from the control library were grouped into 381 UniGene entries whereas only 33.7% (379/1124) of ESTs from the Cd-treated library grouped into 162 UniGene entries.
2. Data suggest that a larger proportion of unique EST sequences (with no homology to any Genbank entry) were discovered in the Cd-treated library than in the control library.
3. It is possible that these unique sequences may relate to TE-induced genome rearrangements.

Transcriptomics study identified pathways associated with Cadmium exposure

- Oxidative phosphorylation
- Ubiquitin pathway
- Stress response
- Translational machinery
- Antimicrobials
- Antivirals - Hemocyanin
- Cytoplasmic genes
- Metal binding
- Cuticle proteins
- Transposable elements-like, and re-arranged sequences**
- ESTs of unknown function

There are differences in TE-like repeats in control and Cd-treated larvae

R2 non-LTR retrotransposon is induced by cadmium

Summary Table	control PL42		Summary Table	Cd-treated PL42	
Repeat Class	**Fragments**	**Length**	**Repeat Class**	**Fragments**	**Length**
Interspersed Repeat	6	532	Interspersed Repeat	2	267
DNA transposon	**174**	**12815**	**DNA transposon**	**104**	**6443**
EnSpm	24	2523	EnSpm	11	1053
Helitron	47	2161	Helitron	40	2049
Kolobok	3	198	Kolobok	1	50
Mariner/Tc1	11	1123	Mariner/Tc1	9	442
MuDR	16	834	MuDR	6	340
P	7	826	P	2	168
Polinton	27	2268	Polinton	4	360
Harbinger	1	84			
Sola	3	329	Sola	3	169
Zator	1	77	Zator	1	34
hAT	16	1160	hAT	12	718
piggyBac	3	214	piggyBac	1	139
			Transib	2	111
Endogenous Retrovirus	**26**	**2122**	**Endogenous Retrovirus**	**4**	**309**
ERV1	13	1286	ERV1	2	194
ERV2	11	758	ERV2	2	115
ERV3	1	46			
LTR Retrotransposon	**59**	**5401**	**LTR Retrotransposon**	**47**	**4436**
BEL	4	522	BEL	4	283
Copia	10	633	Copia	7	530
DIRS	2	410	DIRS	2	139
Gypsy	39	3446	Gypsy	33	3411
Non-LTR Retrotransposon	**58**	**5817**	**Non-LTR Retrotransposon**	**68**	**5027**
CR1	20	2041	CR1	16	1203
Crack	2	170			
L1	17	1731	L1	13	1137
NeSL	1	231	NeSL	1	160
Penelope	3	203	Penelope	2	196
RTE	4	348			
SINE	1	142			
SINE2/tRNA	1	142			
Tx1	5	531	Tx1	1	51
			I	2	288
			Jockey	1	97
			Outcast	1	123
			R2	**28**	**1556**
			Pseudogene	3	273
			rRNA	3	273
Repetitive Element	4	406	Repetitive Element	4	203
Simple Repeat	12	2370	Simple Repeat	8	1625
Satellite	11	1814	Satellite	8	1625
MSAT	8	1457	MSAT	5	1102
SAT	3	357	SAT	3	523

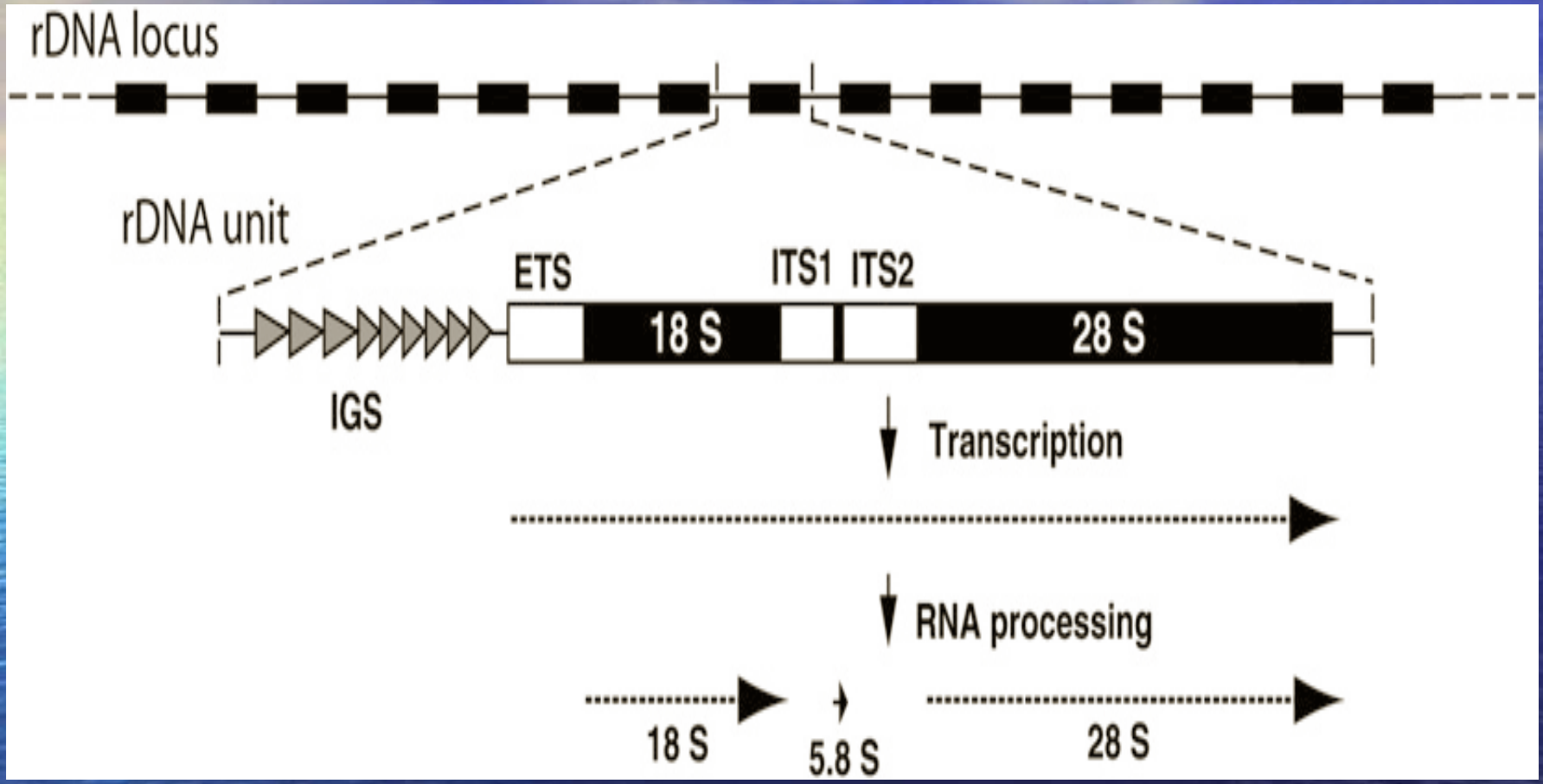
CENSOR analyses of ALL ESTs

1. Transposable Elements (TE)-like repeats varied in Control and Cd-treated PL42

2. increased
    1. Ginger 2 DNA transposon
    2. Copia LTR retrotransposon
    3. Endogenous retrovirus (ERV2)
    4. R2 non-LTR retrotransposon (inserts into 28s rRNA)*

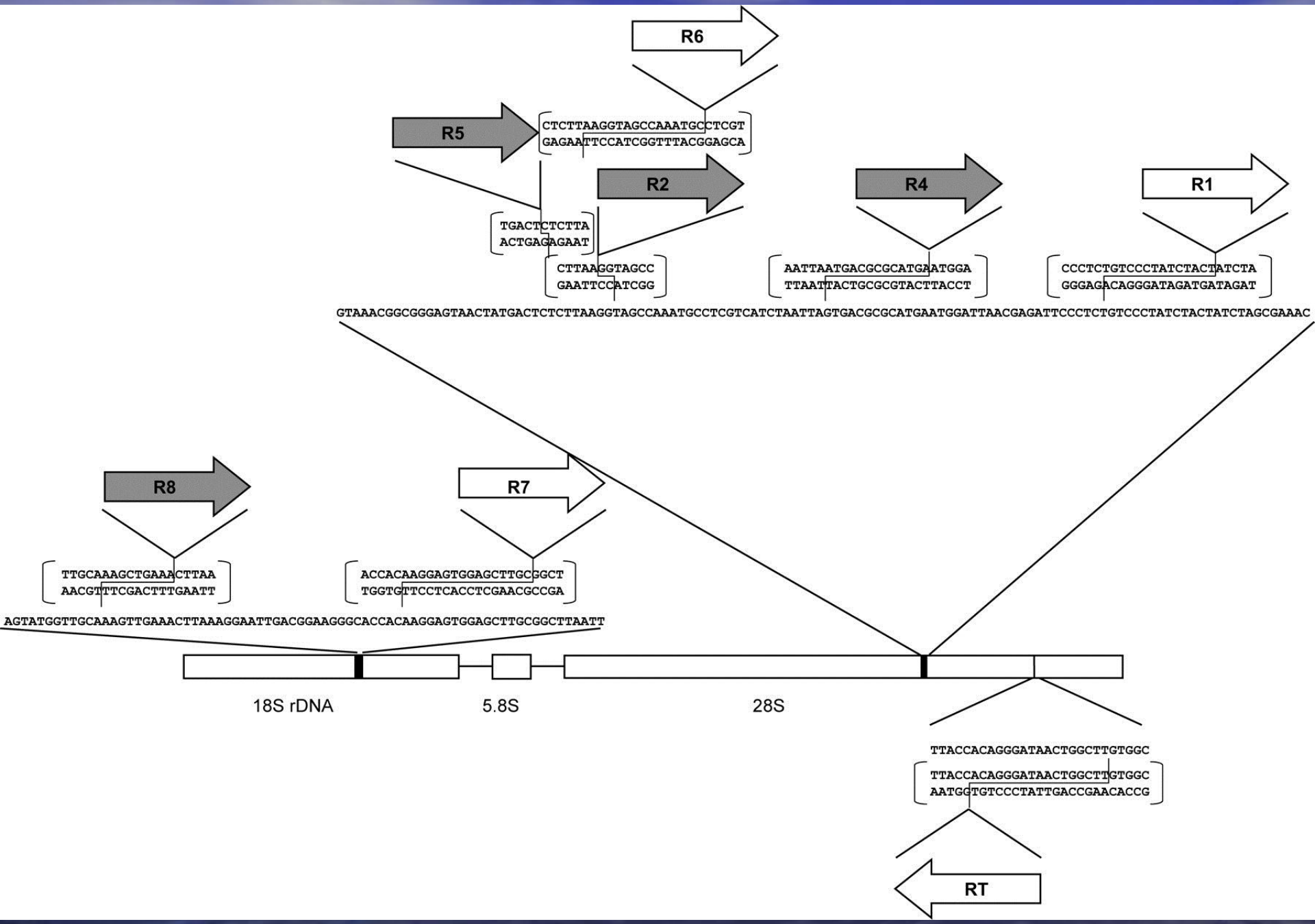
3. decreased
    1. EnSpm, MuDR, Polinton –all except Ginger 2
    2. Gypsy, BEL
    3. L1, CR1,

Organization of the ribosomal RNA (rRNA) genes in eukaryotes



rRNA genes are organized into tandemly repeated units as diagrammed at the top. A typical unit is shown in expanded detail.

Eickbush & Eickbush (2007)


R6
R5
CTCTTAAGGTAGCCAAATGCCTCGT
GAGAATTCCATCGGTTTACGGAGCA
R2
R4
R1
TGACTCTCTTA
ACTGAGAGAAT
CTTAAGGTAGCC
GAATTCCATCGG
AATTAATGACGCGCATGAATGGA
TTAATTACTGCGCGTACTTACCT
CCCTCTGTCCCTATCTACTATCTA
GGGAGACAGGGATAGATGATAGAT
GTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGATTAACGAGATTCCCTCTGTCCCTATCTACTATCTAGCGAAAC
R8
R7
TTGCAAAGCTGAAACTTAA
AACGTTTCGACTTTGAATT
ACCACAAGGAGTGGAGCTTGCGGCT
TGGTGTTCCTCACCTCGAACGCCGA
AGTATGGTTGCAAAGTTGAAACTTAAAGGAATTGACGGAAGGGCACCACAAGGAGTGGAGCTTGCGGCTTAATT
18S rDNA
5.8S
28S
TTACCACAGGGATAACTGGCTTGTGGC
TTACCACAGGGATAACTGGCTTGTGGC
AATGGTGTCCCTATTGACCGAACACCG
RT

Development of EST-SSRs for mapping – work in progress

- Performed an in-depth review of all papers on linkage maps:
- 1 EST-SSR (Su233) that is homologous to the MLC sequences from Unigene cluster Lva. 1302 was mapped to LG7 (Alcivar-Warren et al. 2007)
- 1 SNP for Arginine kinase was reported (Ciobanu et al. 2010), but has not been mapped (Du et al. 2011).
- Work is underway to develop additional allergen EST-SSR markers using the vast number of ESTs available for all 4 allergens.

Summary

- The results provide baseline expression data on the potential effects of Cd on shrimp health and early growth, and suggest a complex interaction between environmental conditions and host genetic background, which should be further investigated under both laboratory and commercial conditions.

Acknowledgements

- FUCOBI Foundation, Guayaquil, Ecuador

- Environmental Genomics Inc., Southborough, MA USA

- Biology Department, Boston College, Chestnut Hill, MA USA