# Investigating the Possible Association of Virus with Orocutaneous Lesions in Brown Bullhead

Submitted to Sean Rafferty, Pennsylvania Sea Grant

Submitted by Luke R. Iwanowicz and Deborah D. Iwanowicz
Fish Health Branch
Leetown Science Center, Kearneysville, WV 25430

## Introduction

The observation of tumors in brown bullheads is currently used as a 'Beneficial Use Impairment (BUI)' in Great Lakes Areas of Concern. An increased incidence of hepatic neoplasms is associated with polyaromatic hydrocarbons (PAHs) and other contaminant exposure in this species (Harshbarger and Clark 1990; Baumann and Harshbarger 1995; Blazer et al 2009b). Likewise skin tumors, including papillomas and squamous cell carcinomas have been used as indicators of chemical exposure in bullhead and other species (Grizzle et al. 1981; Smith et al. 1989; Black and Baumann 1991; Pinkney et al. 2001; Blazer et al 2009a). Although no cause-and-effect correlations have been established for wild populations, papillomas have been experimentally induced in brown bullhead by repeated dosing of the skin with sediment extracts that contained high levels of PAHs (Black et al. 1985). Based on the syntax of this BUI definition there is no distinction between liver tumors and skin tumors. This BUI designation is based on the assumption that contaminant exposure is the sole inducer of tumors in this species. While aryl hydrocarbon receptor ligands such as PAHs have been identified in locations such as Presque Isle Bay of Lake Erie, the association between these ligands and skin tumors is weak (Gray et al. 2003; Rafferty et al. 2009). Additionally, in locations outside the Great Lakes the correlation between the prevalence of liver and skin tumors is spurious as well (Pinkney et al. 2001; Pinkney et al. 2012).

While there is considerable evidence that contaminants are associated with skin tumors in the brown bullhead, biotic factors including viruses are known to induce tumors in mammals and lower vertebrates. Approximately 15-20% of cancers in humans are associated with viruses (McLaughlin-Drubin and Munger

2008). Representatives of such viruses include members of the families *Adenoviridae*, *Circoviridae*, *Flaviviridae*, *Herpesviridae*, *Hepadnaviridae*, *Pappilomaviridae*, *Polyomaviridae* and *Retroviridae*. Viruses of the *Herpesviridae* and *Retroviridae* are associated with skin tumors in fish (reviewed in Getchell et al. 1998). Of note, the possibility of a viral etiology for orocutaneous tumors in bullheads has been investigated previously. Edwards and Samsonoff (1977) reported the presence of intracytoplasmic, virus-like particles via electron microscopic examination of a bullhead papilloma; however, this observation has not been reported since that original investigation (Bowser et al. 1991). Their efforts have been made as well and are summarized in Rafferty et al. (2009). Of note, many viruses that induce skin tumors are uncultureable by standard tissue culture methods. It is possible that previous investigations specific to a putative viral etiology for bullhead skin tumors may have been hindered by this characteristic. Additionally, exploration for viruses via electron microscopy is akin to searching for a needle in a haystack especially if virus titer is low. Nucleic acid-based techniques are now available that allow sequence independent amplification and screening of transcriptomes and genomes making discovery of novel genomes (viral, prokaryotic and eukaryotic) possible. Techniques developed for the specific purpose of novel virus identification are also available. (Hanson et al. 2006, Biagini et al. 2007, Nanda et al 2008). Additionally, the discovery of viruses has been enhanced with Next Generation sequencing technologies (Løvoll et al. 2010, Daly et al. 2011).

**Methods**:

*Sample Collection* – During late May 2010, brown bullheads were collected from Sarah's Cove and Lagoons. Fish were transported to the National Fish Health Laboratory, Kearneysville, WV. Upon arrival all fish were either moribund or dead. Live fish were euthanized with a lethal dose of MS-222 and skin tumors were be excised using a sharp scalpel. Tissues were then transferred to RNA Later and stored for molecular analyses. June 2011, twelve additional bullheads were captured around Presque Isle Bay, PA. Lip tumors were removed, placed into vials,

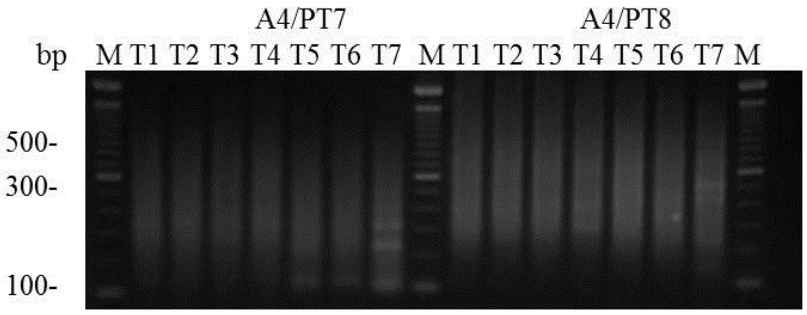
and frozen at -80 C. Frozen vials were then shipped to the National Fish Health Laboratory, Kearneysville, WV where they were stored at -80 C until extraction. These latter samples were used exclusively for the 454 pyrosequencing.

In addition to tumors from bullheads collected in Pennsylvania, early stage orocutaneous neoplasms lesions were procured from an ongoing study in the South River of the Chesapeake Bay watershed. These samples were included in the pyrosequencing sample as well. Historic samples of walleye skin tumors infected with walleye dermal sarcoma virus (WDSV-1) were kindly provided by Dr. Paul Bowser (Cornell University) to serve as a positive control for the PEER and DNA based methods.

*Sample processing* – DNA and RNA were extracted from normal tissue and overt tumors using the DNeasy Tissue Kit (Qiagen) or Total RNA Kit I (Omega Bio Tek) nucleic acid purification kits respectively. RNA samples from bullheads and walleye were reverse transcribed into cDNA for the subsequent primer-extension-enrichment-reaction (PEER) method (Biagani et al 2007). Total RNA was analyzed for quality/ degradation using an Agilent 2100 Bioanalyzer. These methods have been modified in our lab, and have facilitated the successful identification of a fish Herpesvirus (channel catfish virus) and amphibian Iridovirus (frog virus-3) grown in tissue culture and in host tissue (Figure 1).

*Quality Control/ Verification of WDSV-1 infection*- In order to verify that control walleye skin was negative for WDSV-1 and that intact viral nucleic acid was present in the tumor tissue, a simple RT-PCR was conducted. Genomic DNA was used as template from the skin as this represents the available nucleic acids in the driver for the PEERmod. Template from the tumor was cDNA. Both templates were amplified using the primer pair WDSV-1 P3 and WDSV-1 P4 which target a 343 bp amplicon of the WDSV-1 genome (Table 1; Poulet et al. 1996).

**Figure 1.** PEER amplification profile obtained either directly from tissue or BF2 cell culture supernatant inoculated with cell free extract from infected tissues. Final PCR results using A4/PT7 or A4/PT8 primers are shown. M: molecular weight markers. T1: negative control. T2: tumor sample taken from a frozen brown bullhead barbel. T3: tumor sample taken from a frozen brown bullhead barbel. T4: tumor tissue sample taken from a brown bullhead barbel preserved in RNAlater. T5: Unknown sample taken from the coelomic tissue of leopard frogs (*Rana pipiens*). T6: positive control Ranavirus isolate grown on BF2 cells initially inoculated with cell free extract from the American toad (*Bufo americanus*). T7: Known channel catfish virus isolate grown on BF2 cells initially inoculated with cell free extract from brown bullhead.




Amplification conditions were 95 °C for 5 minute for initial denaturation followed by 95 °C for 1 minute, 45 °C for 30 sec and 72 °C for 30 sec for 30 cycles. Extension of the amplicons was completed at 72 °C for 5 minutes, and the final products were chilled to 10 °C. Products were resolved in 1.2% agarose electrophoresis.

*PEER* - Extracted DNA (18ul) from above was mixed with 2.5 µL 10x buffer, 1.0 µL 10mM dNTP solution, and 2.0 µL of each forward and reverse primer at 10 pmol. One *driver* (DNA from control tissue) and two *testers* (cDNA and DNA each from tumors) were then processed using 10 pmol of different primers (Table 1). Samples were incubated at 94°C for 2 min and on ice for 2 min. Then2.5 units of 3'-5' exo-Klenow DNA polymerase (New England Biolabs, Ipswich, MA) was added and the samples were incubated at 37 °C for 1 h. This denaturation-annealing-elongation cycle was repeated and then enzymes were heat inactivated at 75 °C for 10 min. The protocol was then continued as described by Biagini et al. (2007).

*PEERmod* – A modified version of the protocol was tested using only genomic DNA as a driver. This approach in theory would allow for the detection of unique genomic DNA or expressed transcripts between the samples.

**Table 1**. Primer sequences

| **Primer Name** | **Sequence** |
|---|---|
| *Driver* | |
| PDSMART | AAGCAGTGGTAACAACGCAGAGTACGCGGG |
| PDON7 | AAGCAGTGGTAACAACGCAGAGTAIIIIIII |
| *Tester* | |
| PT1G | ACACTCGAGGAGGTCTGGAGGGG |
| PT1N7 | ACACTCGAGGAGGTCTGGAGIIIIIII |
| *Tester II* | |
| PT8G | AAGCAGAGGCAGCATTGGAGGG |
| PT7N7 | GAGCTGTGGTGAGTTGGTTGGAAIIIIIII |
| *Universal* | |
| SP6 | TATTTAGGTGACACTATAG |
| T7 | TAATACGACTCACTATAGGG |
| WDSV-1 P3 | TGAAGCAGGAATACCTACCT |
| WDSV-1 P4 | CTGTAAGTCCGTTCTCTTGT |
| Cons Lower | cccgaattcagatcTCNGTRTCNCCRTA |
| Adeno | gggaattctaGAYATHTGYGGNATGTAYGC |
| HV | cggaattctaGAYTTYGCNWSNYTNTAYCC |
| I = Inosine | |

PEER products were then cloned into the pCR-XL-TOPO vector (Invitrogen) and transfected into Mach1-T1 chemically competent E. coli. Colonies were expanded in S.O.C medium and plated onto LB agar containing 50ug/ ml of kanamycin. The resulting clones were PCR amplified using M13 primers provided with the TOPO-XL cloning kit to verify the presence of inserts.

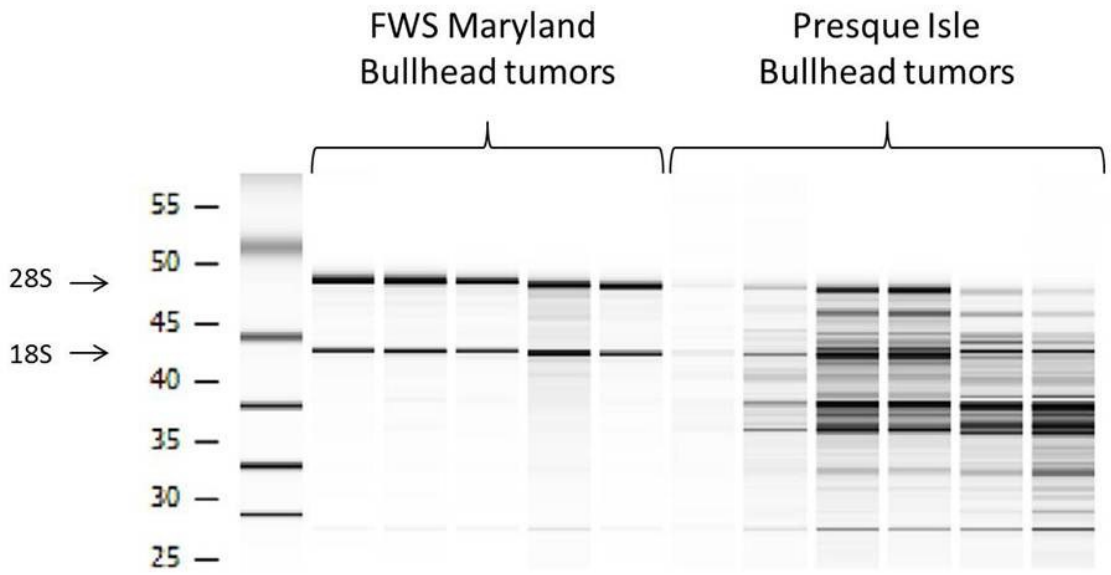
PEERmod products were cloned using a pGEM-T Easy Vector kit (Promega). Positive clones were amplified with SP6 and T7 primers. Clones were then heat lysed and plasmid DNA was harvested for direct sequencing. Sequences were

identified via nucleotide blast (Blastn) or Blastx searches to the NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

*Large DNA virus PCR* – A number of tumorogenic viruses contain a large DNA genome. Hanson et al. (2006) developed a method that allows for the amplification of such genomes using degenerate primer sets that target the highly conserved DNA polymerase gene. Extracted DNA from pooled brown bullhead skin tumors were amplified with the primer pairs Cons Lower and Adeno, or Cons Lower and HV. Primer sequences are listed in Table 1. PCR was conducted in a Rotor-Gene Q using GoTaq qPCR mastermix. PCR used the appropriate forward primer and the consensus reverse primer (Table 1). The reaction conditions were: 93°C, 1 min for one cycle followed by 93°C, 30 sec; 45°C, 2 min; 72°C 3 min for 35 cycles followed by a single cycle at 72°C for 4 min. Product was evaluated by electrophoresis on 1.5% agarose gels. Products were cloned into the using the PGEM cloning kit as described above.

*Pyrosequencing* – Total RNA was extracted following manufacturer's instructions using the E.Z.N.A.™ Total RNA Kit I (Omega Bio-tek, Norcross, GA), from 17 specimens: five brown bullhead lip tumors excised from fish collected from the South River in Maryland (September 2011) and 12 samples from orocutaneous tumors of bullheads inhabiting locations within Presque Isle Bay, PA (June 2011). Total RNA was quantified and analyzed for RNA integrity with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). This analysis identified extensive degradation of sample RNA from the Presque Isle Bay samples (Figure 2). As a result on the 3 best samples from Presque Isle Bay were pooled with the high quality samples from the MD collections. Total RNA samples from 8 fish were pooled at equimolar concentrations for downstream processing.

**Figure 2.** Bioanalyzer trace of total RNA from brown bullhead tissues. Trace indicates significant RNA degradation denoted by the lack of distinct 18S and 28S peaks and an accumulation of low molecular weight fragments.



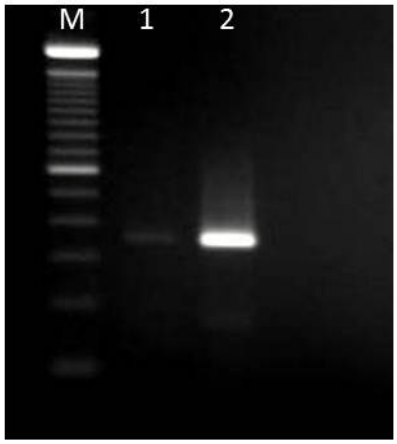

Ribosomal RNA was removed from total RNA using a RiboZero rRNA Removal Kit (Human/Mouse/Rat; Epicenter) according to manufacturer instructions. A cDNA rapid library was then prepared for the GS Junior Titanium Series (June 2010 Rev) using manufacture protocols. The resulting sequences were imported into CLC Workbench for assembly. Blastn and blastx searchers were then conducted against NCBI nt and nr databases. Filters for the keywords 'virus', 'oncogene', and 'interferon' to sort through the lowest e-value gene descriptions.

## Results and Discussion

*Quality Control*

Analysis of extracted RNA from bullhead and walleye tumors (not shown) revealed low RNA quality. This was apparent by the absence of clear 18S and 28S rRNA peaks and an accumulation of small fragments in the Bioanalyzer trace and low RIN values

**Figure 3.** Amplification of the WDSV-1. DNA from control walleye skin (1) and cDNA from a WDSV-1 infected skin tumor.




(Figure 2). Given that these were the only available samples, however, we continued with the virus discovery analysis as proposed.

Amplification of WDSV-1 was successful from the cDNA synthesized from the tumor samples. A weak band of the appropriate size was amplified from control DNA however (Figure 3). This indicates that WDSV-1 had integrated into the host DNA of the control fish, or that some viral replication was in progress in the absence of a tumor. As a result, the negative control was deemed inadequate for the PEER method.

***PEER***

- The PEER generated less than 20 clones for bullhead samples. None of the inserts yielded sequence with homology to viruses in the NCBI database.
- Approximately 100 clones resulted from the PEER for the walleye samples. Similar to the results for the bullhead, none of the sequenced inserts shared homology to viruses in the NCBI database.

***PEERmod***

- The PEERmod generated approximately 100 clones for bullhead samples. None of the inserts yielded sequence with homology to viruses in the NCBI database.
- Approximately 300 clones resulted from the PEER for the walleye samples. Similar to the results for the bullhead, none of the sequenced inserts shared homology to viruses in the NCBI database.

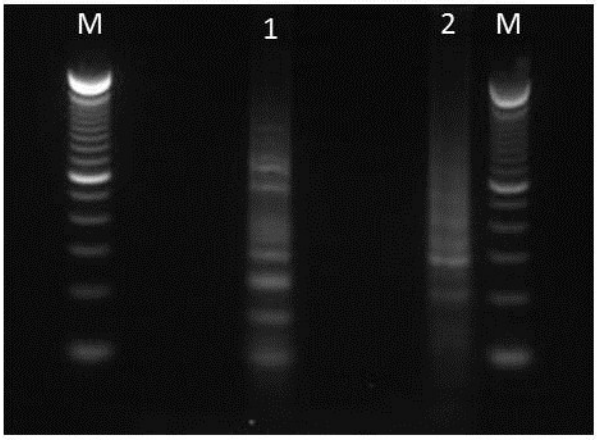
***Large DNA virus PCR***

- Amplification of DNA extracted from brown bullhead skin tumors with the degenerate DNA polymerase resulted in numerous amplicons (Figure 4).
- None of these amplicons had similar homology to viruses in the NCBI database. Interestingly three brown bullhead genes were identified and included 1) toll-like receptor 3, Immunoglobulin heavy chain and the NOD1 receptor.

***Pyrosequencing – Roche 454 Jr.***

The pyrosequencing yielded a total of 98,997 reads with a mean length of 330 bases. Summary statistics and a summary mapping report are contained in Appendix 1. The assembly of these sequences led to the computational construction of 10,093 contigs. Singleton sequences that were not assembled into contigs were represented by 30,329 sequences with a mean length of 310 bases. Approximately 10% of the contigs represented ribosomal RNA suggesting that the mRNA enrichment step was successful. These contigs and raw sequence reads will be deposited into the publicly searchable Sequence Read Archive

**Figure 4.** Amplification of DNA isolated from brown bullhead skin tumors with degenerate primers for Herpesvirus and Adenovirus DNA polymerase. Numerous amplicons generated with the HV primer (1) and Adeno primer (2).




(http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=announcement) maintained by National Center for Biotechnology Information (NCBI) or other depository (Wheeler et al. 2008).

While definitive viral sequences were not identified a number of gene transcripts associated with cellular responses to viral infection were observed (Table 2). Additionally a number of oncogene and transcripts commonly associated with tumors were observed (Table 3). A complete list of putative gene IDs for all contigs are included in Appendix 2.

**Table 2a**. Host genes commonly associated with viral infection (Only those with an E-value > than 1 x $10^{-5}$ are reported). Putative ID based on Blastn.

| Query | Lowest E-value | Accession (E-value) | Putative ID |
|---|---|---|---|
| Ameiurus_contig_779 | 0 | DQ353791 | Interferon gamma inducible protein 30 |
| Ameiurus_contig_7504 | $1.45^{-137}$ | NM_001200299 | Interferon 2 |
| Ameiurus_contig_1273 | $7.21^{-90}$ | FM864346 | Interferon gamma2 |
| Ameiurus_contig_6581 | $2.14^{-81}$ | AM887792 | Virus induced TRIM protein (finTRIM gene) |
| Ameiurus_contig_9431 | $9.01^{-72}$ | NM_001200167 | Interferon induced protein 2 |
| Ameiurus_contig_8852 | $9.56^{-53}$ | XM_678994 | Interferon-induced very large GTPase 1-like |
| Ameiurus_contig_2346 | $2.27^{-49}$ | AM887792 | Virus induced TRIM protein (finTRIM gene) |
| Ameiurus_contig_5240 | $9.73^{-48}$ | AM887815 | Virus induced TRIM protein (finTRIM gene) |
| Ameiurus_contig_1454 | $1.17^{-42}$ | AM887805 | Virus induced TRIM protein (finTRIM gene) |
| Ameiurus_contig_511 | $9.44^{-41}$ | NM_001008614 | Interferon regulatory factor 2 |
| Ameiurus_contig_6758 | $1.31^{-39}$ | XM_693658 | FinTRIM family, member 7 (ftr07) |
| Ameiurus_contig_5561 | $1.01^{-37}$ | NM_001008614 | Interferon regulatory factor 2 |
| Ameiurus_contig_5385 | $5.54^{-27}$ | XM_002660912 | FinTRIM family, member 23 (ftr23) |
| Ameiurus_contig_8993 | $8.51^{-21}$ | XM_003446312 | Interferon-induced very large GTPase 1-like |
| Ameiurus_contig_160 | $2.73^{-19}$ | XM_003197983 | FinTRIM family, member 65 (ftr65) |
| Ameiurus_contig_4516 | $1.51^{-18}$ | FM864346 | Interferon gamma2 |
| Ameiurus_contig_9992 | $1.82^{-14}$ | XM_678994 | Interferon-induced very large GTPase 1-like |
| Ameiurus_contig_7293 | $6.62^{-11}$ | CR545476 | Similar to vertebrate tripartite motif-containing 54 |
| Ameiurus_contig_1818 | $3.45^{-11}$ | DQ353791 | Interferon gamma inducible protein 30 |
| Ameiurus_contig_3343 | $3.015^{-10}$ | XM_678037 | Interferon-induced, hepatitis C-associated microtubular aggregate |
| Ameiurus_contig_8937 | $5.88^{-09}$ | BX465210 | Novel protein similar to interferon gamma 1 |
| Ameiurus_contig_2367 | $3.49^{-08}$ | XM_003451121 | E3 ubiquitin/ISG15 ligase TRIM25-like |
| Ameiurus_contig_7598 | $1.64^{-08}$ | XM_003460109 | Interferon-induced very large GTPase 1-like |
| Ameiurus_contig_9581 | $4.47^{-06}$ | GU725006 | Interferon-inducible 58 kDa protein |
| Ameiurus_contig_1538 | $8.83^{-05}$ | XM_003199710 | E3 ubiquitin-protein ligase TRIM8-like |

**Table 2b**. Host genes commonly associated with viral infection (Only those with an E-value > than 1 x $10^{-5}$ are reported). Putative ID based on Blastx.

| Query | Lowest E-value | Accession (E-value) | Putative ID |
|---|---|---|---|
| Amieurus_contig_6477 | $9.00^{-99}$ | NP_001187103 | Interferon-induced GTP-binding protein Mx1 |
| Amieurus_contig_9564 | $7.55^{-86}$ | CAP08949 | Fish virus induced TRIM protein |
| Amieurus_contig_779 | $1.6^{-79}$ | NP_001006057 | Interferon gamma inducible protein 30 |
| Amieurus_contig_1538 | $6.02^{-69}$ | XP_003199758 | Probable E3 ubiquitin-protein ligase TRIM8-like |
| Amieurus_contig_5240 | $3.65^{-65}$ | CAP08948 | Fish virus induced TRIM protein |
| Amieurus_contig_8852 | $2.27^{-61}$ | XP_684086 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_7790 | $2.78^{-59}$ | CBN80878 | Tripartite motif-containing protein 62 |
| Amieurus_contig_6581 | $4.34^{-57}$ | ACO13467 | Tripartite motif-containing protein 25 |
| Amieurus_contig_8993 | $1.21^{-56}$ | XP_684086 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_5258 | $1.34^{-54}$ | AAP49829 | Interferon-inducible protein Gig2 |
| Amieurus_contig_4073 | $9.49^{-53}$ | CAP08949 | Fish virus induced TRIM protein |
| Amieurus_contig_9581 | $9.81^{-52}$ | AAP42145 | Interferon-inducible protein IFI58 |
| Amieurus_contig_9805 | $8.18^{-49}$ | CAP08959 | Fish virus induced TRIM protein |
| Amieurus_contig_7345 | $5.93^{-48}$ | XP_001346132 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_5313 | $7.24^{-45}$ | NP_001002870 | Trim24 protein |
| Amieurus_contig_3893 | $2.7^{-44}$ | NP_001038633 | Novel protein similar to vertebrate tripartite motif |
| Amieurus_contig_5385 | $4.06^{-43}$ | NP_001038735 | finTRIM family, member 14 |
| Amieurus_contig_160 | $2.85^{-40}$ | NP_001038735 | finTRIM family, member 14 |
| Amieurus_contig_3988 | $8.33^{-40}$ | CAQ14367 | finTRIM family protein |
| Amieurus_contig_3512 | $1.84^{-39}$ | XP_003198092 | Interferon-induced very large GTPase 1 |
| Amieurus_contig_9554 | $3.55^{-38}$ | NP_001187103 | Interferon-induced GTP-binding protein Mx1 |
| Amieurus_contig_7504 | $2.61^{-37}$ | NP_001187228 | Interferon 2 |
| Amieurus_contig_158 | $2.63^{-37}$ | NP_001038735 | finTRIM family, member 14 |
| Amieurus_contig_7598 | $1.64^{-36}$ | XP_684086 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_4457 | $2.16^{-35}$ | XP_003201305 | E3 ubiquitin-protein ligase TRIM21 |
| Amieurus_contig_2346 | $6.4^{-32}$ | CAP08960 | Fish virus induced TRIM protein |
| Amieurus_contig_3488 | $3.23^{-31}$ | AAP49828 | Interferon-inducible protein Gig1 |
| Amieurus_contig_3791 | $9.3^{-31}$ | CAQ14367 | finTRIM family protein |
| Amieurus_contig_1358 | $3.84^{-29}$ | CAP08960 | Fish virus induced TRIM protein |
| Amieurus_contig_1454 | $1.29^{-25}$ | CAQ14348 | finTRIM family protein |
| Amieurus_contig_6758 | $5.09^{-21}$ | CAP08952 | Fish virus induced TRIM protein |
| Amieurus_contig_4562 | $8.66^{-21}$ | XP_002939897 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_2947 | $4.46^{-17}$ | XP_700104 | E3 ubiquitin/ISG15 ligase TRIM25 isoform 3 |
| Amieurus_contig_9992 | $7.82^{-17}$ | XP_684086 | Interferon-induced very large GTPase 1-like |
| Amieurus_contig_9108 | $4.3^{-14}$ | XP_003200077 | Tripartite motif-containing protein 39-like |
| Amieurus_contig_8724 | $4.76^{-14}$ | XP_003197665 | Interferon-induced protein 44-like |
| Amieurus_contig_8825 | $1.03^{-05}$ | XP_700104 | E3 ubiquitin/ISG15 ligase TRIM25 isoform 3 |

**Table 3a**. Oncogenes and tumor associated genes expressed in bullhead orocutaneous tumors (Only those with an E-value > than 1 x 10-5 are reported). Putative ID based on Blastn.

| Query | Lowest E-value | Accession (E-value) | Putative ID |
|---|---|---|---|
| Ameiurus_contig_6656 | 0 | GU587872 | Tumor protein d54 (TPD54) |
| Ameiurus_contig_7029 | 0 | GU588158 | Translationally-controlled tumor protein-like protein |
| Ameiurus_contig_4375 | 0 | NM_001200465 | RAS oncogene (rab3-like) |
| Ameiurus_contig_3950 | $1.4^{-171}$ | GU588158 | Translationally-controlled tumor protein-like protein |
| Ameiurus_contig_5650 | $2.32^{-169}$ | NM_201454 | RAS oncogene family (rab2a) |
| Ameiurus_contig_6100 | $2.2^{-164}$ | NM_001200076 | Tumor protein p53 (tp53) |
| Ameiurus_contig_6595 | $3.1^{-135}$ | NM_001200076 | Tumor protein p53 (tp53) |
| Ameiurus_contig_4609 | $1.95^{-135}$ | NM_200928 | RAS oncogene family (rab7) |
| Ameiurus_contig_2634 | $9.78^{-133}$ | AY326792 | P55 c-fos oncogene |
| Ameiurus_contig_5466 | $3.3^{-121}$ | BC165809 | Tumor protein p63 |
| Ameiurus_contig_5075 | $1.84^{-119}$ | NM_001122705 | V-abl Abelson murine leukemia viral oncogene homolog 2 |
| Ameiurus_contig_946 | $2.01^{-100}$ | BX005210 | Braf gene for v-raf murine sarcoma viral oncogene |
| Ameiurus_contig_8805 | $4.84^{-89}$ | BC062857 | RAS oncogene family (rab1a) |
| Ameiurus_contig_9226 | $5.67^{-89}$ | NM_001140089 | V-ets erythroblastosis virus E26 oncogene homolog 2 |
| Ameiurus_contig_9252 | $5.69^{-89}$ | AY423018 | RAS oncogene family (RAP1B) |
| Ameiurus_contig_5875 | $1.40^{-73}$ | NM_001003449 | RAS oncogene family (rab18b) |
| Ameiurus_contig_7955 | $2.15^{-72}$ | NM_001200574 | Receptor-binding cancer antigen expressed on siso cells |
| Ameiurus_contig_1771 | $1.4^{-64}$ | GU588158 | Translationally-controlled tumor protein-like protein |
| Ameiurus_contig_995 | $6.31^{-62}$ | NM_201329 | RAS oncogene family (rab13) |
| Ameiurus_contig_1495 | $6.93^{-61}$ | NM_001004597 | Danio rerio disrupted in renal carcinoma 2 |
| Ameiurus_contig_2733 | $2.18^{-59}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_4544 | $7.43^{-59}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_4467 | $1.53^{-58}$ | NM_001002089 | Tumor susceptibility gene 101 |
| Ameiurus_contig_6090 | $1.03^{-53}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_8562 | $2.66^{-50}$ | AF299060 | Squamous-cell carcinoma T-cell-recognized antigen |
| Ameiurus_contig_7638 | $1.33^{-45}$ | DQ086208 | N-ras oncogene p21 mRNA |
| Ameiurus_contig_2050 | $2.55^{-44}$ | NM_001013270 | Yamaguchi sarcoma viral oncogene homolog 1 |
| Ameiurus_contig_1933 | $8.60^{-44}$ | NM_001140681 | V-ral simian leukemia viral oncogene homolog Bb |
| Ameiurus_contig_7038 | $1.25^{-41}$ | NM_001007312 | FBJ murine osteosarcoma viral oncogene homolog B |
| Ameiurus_contig_9284 | $1.62^{-40}$ | XM_003444218 | Tumor protein 63-like, transcript variant 2 |
| Ameiurus_contig_10081 | $6.59^{-39}$ | NM_199769 | Suppression of tumorigenicity 13 |
| Ameiurus_contig_5795 | $5.3^{-39}$ | NM_199853 | Wilms tumor 1 associated protein (wtap) |
| Ameiurus_contig_79 | $1.10^{-37}$ | XM_003454260 | Ski oncogene-like |
| Ameiurus_contig_5878 | $1.21^{-36}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_5689 | $1.4^{-32}$ | XM_003444217 | Tumor protein 63-like, transcript variant 1 |
| Ameiurus_contig_8946 | $7.66^{-26}$ | XM_002666744 | Lethal(3)malignant brain tumor-like protein 3-like |

| | | | |
|---|---|---|---|
| Ameiurus_contig_4086 | $1.21^{-25}$ | BX510928 | RAS oncogene family |
| Ameiurus_contig_640 | $3.32^{-21}$ | NM_001173687 | Suppression of tumorigenicity 5 |
| Ameiurus_contig_3168 | $2.23^{-20}$ | NM_201329 | RAS oncogene family (rab13) |
| Ameiurus_contig_5997 | $6.66^{-18}$ | NM_001002673 | Squamous cell carcinoma antigen |
| Ameiurus_contig_2373 | $1.91^{-15}$ | NM_001111156 | Suppression of tumorigenicity 7 like |
| Ameiurus_contig_2648 | $3.53^{-15}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_1561 | $5.74^{-14}$ | BC128654 | Danio rerio suppression of tumorigenicity 14 |
| Ameiurus_contig_5971 | $1.07^{-13}$ | XM_001920023 | FAT tumor suppressor homolog 2 |
| Ameiurus_contig_4246 | $1.45^{-12}$ | BT075067 | Tumor suppressor candidate 2 |
| Ameiurus_contig_3110 | $6.31^{-11}$ | XM_001373285 | Leucine zipper, putative tumor suppressor 1 (LZTS1) |
| Ameiurus_contig_4224 | $8.21^{-10}$ | BC134114 | Danio rerio suppression of tumorigenicity 14 |
| Ameiurus_contig_7905 | $1^{-09}$ | XM_002666744 | Lethal(3)malignant brain tumor-like protein 3-like |
| Ameiurus_contig_2949 | $7.51^{-08}$ | XM_002662510 | Ubiquitin specific protease 4 (proto-oncogene) |
| Ameiurus_contig_812 | $6.27^{-08}$ | NM_001023580 | V-ets erythroblastosis virus E26 oncogene homolog 2 |
| Ameiurus_contig_4484 | $3.97^{-06}$ | BC163970 | Suppression of tumorigenicity 5 |
| Ameiurus_contig_312 | $6.26^{-06}$ | NM_199582 | Tumor protein D52-like 2b (tpd52l2b) |
| Ameiurus_contig_2737 | $4.7^{-05}$ | XM_001920023 | FAT tumor suppressor homolog 2 |

**Table 3b**. Oncogenes and tumor associated genes expressed in bullhead orocutaneous tumors (Only those with an E-value > than 1 x 10-5 are reported). Putative ID based on Blastx.

| Query | Lowest E-value | Accession (E-value) | Putative ID |
|---|---|---|---|
| Amieurus_contig_6475 | $2.5^{-113}$ | XP_002200623 | Similar to v-akt murine thymoma viral oncogene homolog 1 |
| Amieurus_contig_5650 | $1.51^{-108}$ | NP_958862 | Ras-related protein Rab-2A |
| Amieurus_contig_2669 | $2.63^{-108}$ | NP_956417 | RAB11a, member RAS oncogene family, like |
| Amieurus_contig_5466 | $4.64^{-94}$ | AAI52688 | Tumor protein p63 |
| Amieurus_contig_1495 | $9.51^{-93}$ | NP_001004597 | Disrupted in renal carcinoma 2 |
| Amieurus_contig_6871 | $1.27^{-92}$ | NP_001002178 | Ras-related protein rab-7-like |
| Amieurus_contig_4467 | $2.81^{-91}$ | NP_001002089 | Tumor susceptibility gene 101 protein |
| Amieurus_contig_7029 | $1.89^{-90}$ | ADO28114 | Translationally-controlled tumor protein-like protein |
| Amieurus_contig_2787 | $3.56^{-86}$ | CAP09309 | Jun B proto-oncogene |
| Amieurus_contig_8551 | $2.55^{-85}$ | XP_001510596 | Similar to serine/threonine kinase TAO1 |
| Amieurus_contig_7707 | $5.66^{-84}$ | NP_001071205 | Suppression of tumorigenicity 14 |
| Amieurus_contig_5875 | $8.23^{-80}$ | NP_001003449 | Ras-related protein Rab-18-B |
| Amieurus_contig_2518 | $1.21^{-75}$ | NP_001118094 | Transcription factor Myc |
| Amieurus_contig_6100 | $6.40^{-64}$ | NP_001187005 | Cellular tumor antigen p53 |
| Amieurus_contig_8562 | $1.18^{-62}$ | AAG28762 | Squamous-cell carcinoma T-cell-recognized antigen |
| Amieurus_contig_675 | $9.84^{-62}$ | NP_956021 | Similar to human tumor differentially expressed 1 |
| Amieurus_contig_6656 | $1.73^{-60}$ | ADO27830 | Tumor protein d54 |
| Amieurus_contig_570 | $7.97^{-59}$ | AAI52688 | Tumor protein p63 |
| Amieurus_contig_995 | $4.18^{-56}$ | NP_958486 | Ras-related protein Rab-13 |
| Amieurus_contig_1551 | $4.53^{-54}$ | NP_571882 | Bcl2l protein |
| Amieurus_contig_7793 | $1.42^{-51}$ | NP_001187995 | Tumor susceptibility gene 101 protein |

| | | | |
|---|---|---|---|
| Amieurus_contig_8946 | $1.62^{-51}$ | XP_002666790 | Lethal(3)malignant brain tumor-like protein 3-like |
| Amieurus_contig_2610 | $1.65^{-50}$ | ACR83585 | Jun B proto-oncogene |
| Amieurus_contig_7860 | $1.08^{-47}$ | NP_001003628 | Adapter molecule crk |
| Amieurus_contig_1926 | $1.12^{-47}$ | NP_001025294 | Nucleoprotein TPR |
| Amieurus_contig_8754 | $6.49^{-47}$ | XP_001510596 | Similar to serine/threonine kinase TAO1 |
| Amieurus_contig_5902 | $8.31^{-47}$ | CAN88705 | Suppression of tumorigenicity 14 |
| Amieurus_contig_9888 | $9.38^{-44}$ | AAI63200 | Bone morphogenetic protein receptor, type II b |
| Amieurus_contig_9932 | $2.35^{-42}$ | NP_998621 | Serine/threonine-protein kinase 38 |
| Amieurus_contig_7905 | $1.32^{-40}$ | XP_003204259 | Lethal(3)malignant brain tumor-like protein 3-like |
| Amieurus_contig_6923 | $8.05^{-40}$ | NP_997915 | Jun B proto-oncogene, like |
| Amieurus_contig_9284 | $1.00^{-38}$ | AAI52688 | Tumor protein p63 |
| Amieurus_contig_9758 | $8.87^{-38}$ | NP_955966 | Serine/threonine-protein kinase 3 |
| Amieurus_contig_7038 | $8.00^{E-38}$ | CAK04207 | Novel protein similar to vertebrate FBJ murine osteosarcoma viral oncogene homolog B |
| Amieurus_contig_2050 | $1.47^{-36}$ | NP_001013288 | V-yes-1 Yamaguchi sarcoma viral oncogene homolog 1 |
| Amieurus_contig_123 | $1.55^{-33}$ | CAQ14160 | MAP kinase-interacting serine/threonine kinase 2 |
| Amieurus_contig_236 | $6.31^{-32}$ | CAQ15160 | Tumor protein p73 |
| Amieurus_contig_4224 | $9.84^{-32}$ | XP_001112126 | Suppressor of tumorigenicity 14 protein-like |
| Amieurus_contig_8746 | $5.23^{-31}$ | NP_001122291 | V-erb-b2 erythroblastic leukemia viral oncogene homolog 3b |
| Amieurus_contig_8127 | $1.79^{-30}$ | NP_001186798 | Feline sarcoma oncogene |
| Amieurus_contig_1969 | $4.61^{-30}$ | CAX13164 | Novel protein similar to vertebrate translocated promoter region |
| Amieurus_contig_2373 | $2.31^{-29}$ | CBN80808 | Suppressor of tumorigenicity protein 7 homolog |
| Amieurus_contig_8453 | $1.48^{-28}$ | NP_001116177 | Tyrosine-protein kinase ABL2 |
| Amieurus_contig_122 | $1.17^{-25}$ | CAQ14160 | MAP kinase-interacting serine/threonine kinase 2 |
| Amieurus_contig_3110 | $2.01^{-25}$ | XP_697750 | Leucine zipper putative tumor suppressor 1 |
| Amieurus_contig_6595 | $3.01^{-24}$ | NP_001187005 | Cellular tumor antigen p53 |
| Amieurus_contig_742 | $6.07^{-24}$ | NP_001167158 | Suppression of tumorigenicity 5 protein |
| Amieurus_contig_5997 | $2.06^{-23}$ | NP_001002673 | Squamous cell carcinoma antigen recognized by T cells |
| Amieurus_contig_1561 | $2.41^{-23}$ | CAN88704 | Suppression of tumorigenicity 14 |
| Amieurus_contig_9705 | $9.29^{-23}$ | NP_937789 | RAC-beta serine/threonine-protein kinase |
| Amieurus_contig_966 | $8.05^{-21}$ | NP_001164290 | Fas-activated serine/threonine kinase |
| Amieurus_contig_3354 | $2.50^{-20}$ | NP_956063 | Suppression of tumorigenicity 13 |

## Discussion

Here a number of molecular biology techniques were exploited in an attempt to identify the presence of putative viral nucleic acids associated with orocutaneous tumors of the brown bullhead. At the conception of this work Next Generation Sequencing (NGS) was a cost prohibitive option; however, given the rapid development of this technology, a single run of pooled samples became a cost efficient option. While the PEER method has been a useful approach to identify unknown viruses, it is simply no longer cost/ time effective to sequence hundreds to

thousands of clones via Sanger sequencing. However, adaptation could be made to hybridize these approaches. The 454 pyrosequencing platform allowed for the tentative identification of over 10K genes (with some redundancy), and over 30K non-overlapping sequence directly from tumor tissue without *a priori* knowledge of brown bullhead gene sequences. Given the vast sequence data generated by the 454 pyrosequencing the following will focus exclusively on these results.

Here, unequivocal viral nucleic acid was not identified. A few transcripts identified did share sequence homology with an RNA-dependent DNA polymerase and proteins associated with herpeseviruses based on lowest assigned E-values from Blastx searches of the NCBI non-redundant protein database. Based on a more comprehensive examination of the Blast results, the RNA-dependent DNA polymerase is more likely a vertebrate reverse transcriptase. The only predicted herpesvirus protein with an E-value greater than $1*10^{-4}$ also shared high homology to a hemagglutination activity domain protein not commonly observed in vertebrates. Results from the Blastn search, however, lend stronger support (E-value = $1.65*10^{-36}$) for the identity of this sequence as a noncoding region transcript for an inhibitor of apoptosis gene.

While definitive viral nucleic acids were not identified in the current discovery screening, a number of host gene transcripts commonly associated with virus infection were detected (Table 2). These included Interferon-γ, interferon inducible genes and a number of fish tripartite motif- containing (finTRIM) proteins. Interferon-γ is best recognized as an inducible gene associated with host defense to viral infection and inflammation (Farrar and Schreiber 1993). In some tissues, however, it is sometime expressed in response to disease status such as chronic liver disease (Nagao et al 2000). It is typically not expressed in tumors, but is commonly used as an anti-tumor chemotherapeutic agent. While many of the finTRIM proteins are up-regulated in response to viral infection, many are constitutively expressed (van der Aa et al 2012). Thus, while expression of these transcripts is not a specific biomarker of viral infection, the presence of such

transcripts does add to the weight of evidence supporting a virally induced tumor. However, their presence could simply reflect inflammation that is commonly associated with necrotic tumors. Histology was not conducted to verify tumor type or identify inflammation given that these were frozen tissue samples.

Perhaps the most interesting observation in the current investigation was the identification of a number of oncogenes, tumor suppressor genes and those associated with cancer. They are present in many tumor causing viruses, but they are also present in the vertebrate genome. Many of these genes are constitutively expressed and are involved in general cell maintenance and application, and are therefore associated with tumors following mutagenic events or oncogenic viral infection. The development of neoplasia is a complex process and is simply the result of a disrupted homeostatic balance of cell growth and cell death. An oncogene is a gene that has potential to cause cancer. The sequencing of tumors genomes and transcriptomes is now common practice in human medicine (Levin et al. 2009, Morrissy et al 2009, Nicholas et al. 2011). The Cancer Genome Project (http://www.sanger.ac.uk/genetics/CGP/Census/) has identified over 450 genes commonly associated with human cancers implicated via mutation in cancer. A list of these genes is included in Appendix 3. Of these, approximately 90% have somatic mutations in cancer, 20% bear germ line mutations that predispose to cancer and 10% show both somatic and germ line mutations. (Futreal et al. 2004). Homologues of many of these genes were identified in the current discovery transcriptome project. While the complete genome of a normal individual is not known, it may be possible to compare transcriptomes of individual fish to identify such mutations in expressed genes. This would facilitate the converse approach of identifying a viral cause and rule in or discount the possibility of chemically induced tumors.

At the present time an unequivocal viral nucleic acid has not been identified in brown bullhead orocutaneous tumors. Failure to identify a microbial agent, however, does not confirm that these skin tumors are not virally induced; however, there is no evidence to support that viruses are the causative agent. As mentioned

above this sequencing strategy was confined to a number of assumptions. Also, there were concerns with the quality of orocutaneous tumor tissue. The low RNA integrity values observed during the quality control process indicated significant RNA degradation. As a result viral RNA transcripts may have been degraded if present. It is not clear if the poor tissue quality was an artifact of sample preservation or the intrinsic condition of the neoplastic tissue. Parallel histopathology was not conducted on these tissues. The high quality samples included in this work were from of orocutaneous swellings (confirmation as neoplastic is pending) from bullheads collected in the South River of the Chesapeake Bay watershed. Thus these findings, or lack thereof, do not necessarily represent bullhead skin tumors of bullhead from the Great Lakes. Also, fresh, high quality samples would likely increase the chances of identifying a microbial agent if one is present. Additionally, virally induced tumors are typically seasonal in occurrence (Getchell et al. 1998). Sampling bullhead skin tumors from fish collected during multiple seasons may increase the likelihood of capturing the window of peak viral replication. Without accounting for such possible temporal differences, the confidence that a virus is not associated is lessened.

The strategy employed here is not without limitations. The major assumption was that if a virus was present that it was in a replicative phase and that RNA transcripts would be present. Our strategy would not identify the presence of DNA viruses that were not replicating. This approach also would not identify viral nucleic acid integrated into the host genome (a hallmark mechanism of retroviral and herpesvirus biology) unless there was active transcription of those genes. Additionally, we would unlikely identify low copy numbers of non-replicating RNA viruses given that a normalization of total transcripts was not performed. It must also be noted that viruses are usually not complete carcinogens and many years may pass before the manifestation of a tumor. Likewise, depending on the virus, replication does not always occur in tumor cells and the genome is infrequently retained (Butel 2000).

To our knowledge this is the first time high-throughput NGS has been used to investigate the transcriptome of orocutaneous lesions in the brown bullhead. This work has identified a number of putative markers of viral infection and tumorogenesis. More broadly, it has identified over 10K nucleotide sequences that can be used to develop tumor markers or evaluate gene expression in future studies. Given the possibility that gene mutations are an underlying reason for the tumors, this database can be more exhaustively screened for mutations that may lead to protein coding changes. It will also serve as a reference for comparative mutation research. The current sequencing strategy employed here was devised as a simple discovery tool. More quantitative methods (i.e. qPCR) and larger sample sizes will be necessary to document validate tumor biomarker genes. For instance, the overexpression of serine/threonine kinases is associated with tumors. Tumor suppressor genes proteins are well established targets of inactivation for DNA viruses. Moreover, chemically induced mutations in these genes could result in dysregulated cell proliferation and growth.

Given that viruses were not identified here, it is prudent to continue the assumption that contaminant exposure or a genetic predisposition is the cause of these tumors. This possibility of a multi-step combination of virus and contaminant also would not be unprecedented. A classic example is that of the cottontail rabbit papilloma virus. In short, when virus-induced papillomas were treated with coal tar a high percentage of the papillomas converted to carcinomas (Howley and Munger 1999). If indeed these lesions are the result of a transforming virus that integrates into the host genome a more labor intensive full genome sequencing approach may be necessary to confirm viral transformation.

**Acknowledgements**

We would like to thank Sean Rafferty and Jim Grazio for providing fish for this work. We are also grateful to Fred Pinkney of the US Fish & Wildlife Service for providing lesions from brown bullheads from the South River. We would also like to thank Tim King for the use of his sequencing laboratory and analytical software, and Barb Lubinski for her technical assistance.

**References**


Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A and Andersson B. 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proceedings of the National Academy of Sciences **102**: 12891-12896.

Baumann PC and Harshbarger JC. 1995. Decline in liver neoplasms in wild brown bullhead catfish after coking plant closes and environmental PAHs plummet. Environ Health Perspect **103**: 168–170.

Biagini P, de Lamballerie X and de Micco P. 2007. Effective detection of highly divergent viral genomes in infected cell lines using a new subtraction strategy (primer extension enrichment reaction – PEER). J Virol Meth **139**: 106-110.

Black JJ Fox H, Black P and Bock F. 1985. Carcinogenic effects of river sediment extracts in fish and mice. In: Jolley, R.L., R.J. Bull, W.P. Davis, S. Katz, M.H. Roberts, Jr. and V.A. Jacobs (eds). Water Chlorination: Chemistry, Environmental Impact and Health Effects, Vol. 5. Lewis Publishers, Boca Raton, FL, **pp**. 415-428.

Black JJ and Baumann PC. 1991. Carcinogenesis and cancers in freshwater fishes. Environ Health Perspect **90**: 27-33.

Blazer VS, Rafferty SD, Baumann PC, Smith SB and Obert EC. 2009a. Assessment of the "fish tumor or other deformities" beneficial use impairment in brown bullhead (*Ameiurus nebulosus*): I. Orocutaneous tumors. J Great Lakes Res **35**: 517-526.

Blazer VS, Rafferty SD, Baumann PC, Smith SB and Obert EC. 2009b. Assessment of the "fish tumor or other deformities" beneficial use impairment in brown bullhead (*Ameiurus nebulosus*): II. Liver neoplasia. J Great Lakes Res **35**: 527-537.

Bowser PR, Wolfe MJ, Reimer J and Shane BS. 1991. Epizootic Epizootic papillomas in brown bullheads *Ictalurus nebulosus* from Silver Stream Reservoir, New York. Dis Aquat Org **11**: 117–27l.

Butel JS. 2000. Viral carcinogenesis: revelation of molecular mechanisms and etiology of human disease. Carcinogenesis **21**: 405-426.

Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, Quan PL, Briese T, Hornig M, Geiser DM, Martinson V, VanEngelsdorp D, Kalkstein AL, Drysdale A, Hui J, Zhai JH, Cui LW, Hutchison SK, Simons JF, Egholm M, Pettis

JS and Lipkin WI. 2007. A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* **318**: 283–287.

Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, Kellam P, Drou N, Caccamo M, Tiley L, Alexander GJM, Bernal W and Heeney JL. 2011. A Viral Discovery Methodology for Clinical Biopsy Samples Utilising Massively Parallel Next Generation Sequencing. PLOS One **6**: e28879

Edwards MR and Samsonoff WA. 1977. Electron microscopic observations of virus-like particles of a catfish papilloma. Proc A Meet Electron Microsc Soc Am **35**: 394-395.

Farrar MA and Schreiber RD. 1993. The molecular cell biology of interferon-gamma and its receptor. Annu Rev Immunol **11**: 571-611.

Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N and Stratton MR. 2004. A census of human cancer genes. Nature Reviews Cancer **4**: 177-183.

Getchell, RG, Casey JW and Bowser PR. 1998. Seasonal occurrence of virally induced skin tumors in wild fish. J Aquat Ani Health **10**: 191-201.

Gray JP, Leas TL, Obert E, Brown D, Clark GC and Vanden Heuvel JP. 2003. Evidence of arylhydrocarbon receptor ligands in Presque Isle Bay of Lake Erie. Aquat Toxicol **64**: 343-358.

Grizzle JM, Schwedler TE and Scott AL. 1981. Papillomas of black bullheads, *Ictalurus melas* (Rafinesque) living in a chlorinated sewage pond. J Fish Dis **4**: 345-351.

Hanson LA, Rudis MR, Vasquez-Lee M and Montgomery RD. 2006. A broadly applicable method to characterize large DNA viruses and adenoviruses based on the DNA polymerase gene. Virol J **3**: 28-38.

Harshbarger JC and Clark JB. 1990. Epizootiology of neoplasms in bony fish of North America. Sci Total Environ **1**: 1–32.

Howley PM and K Munger. 1999. Human papilloma viruses and squamous cell carcinomas. In Parsoneete, J. (ed) Microbes and Malignancy: Infection as a cause of Human Cancers. Oxford Press, Oxford, UK. **pp** 157-179.

Levin JZ, Berger MF, Adiconis X, Rogov P, Melnikov A, Fennell T, Nusbaum C, Garraway LA and Gnirke A. 2009. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. Genome Biol **10**: R115.

Løvoll M, Wiik-Nielsen J, Grove S, Wiik-Nielsen CR, Kristoffersen AB, Faller R, Poppe T, Jung J, Pedamallu CS, Nederbragt AJ, Meyerson M, Rimstad E, Tengs T. 2010. A novel totivirus and piscine reovirus (PRV) in Atlantic salmon (*Salmo salar*) with cardiomyopathy syndrome (CMS). Virol J **7**: 309.

Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao Y, Hirst M and Marra MA. 2009. Next-generation tag sequencing for cancer gene expression profiling. Genome Res **19**: 1825-35.

Nagao M, Nakajima Y, Kanehiro H, Hisanaga M, Aomatsu Y, Ko S, Tatekawa Y, Ikeda N, Kanokogi H, Urizono Y, Kobayashi T, Shibaji T, Kanamura T, Ogawa S, Nakano H. 2000. The impact of interferon gamma receptor expression on the mechanism of escape from host immune surveillance in hepatocellular carcinoma. Hepat **32**: 491-500.

Nakamura S, Yang CS, Sakon N, Ueda M, Tougan T, Yamashita A, Goto N, Takahashi K, Yasunaga T, Ikuta K, Mizutani T, Okamoto Y., Tagami M, Morita R, Maeda N, Kawai J, Hayashizaki Y, Nagai Y, Horii T, Iida T and Nakaya T. 2009. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. *PLoS ONE*, **4**: e4219.

Nanda S, Jayan G, Voulgaropoulou F, Sierra-Honigmann AM, Uhlenhaut C, McWatters BJP and Patel A. 2008. Universal virus detection by degenerate-oligonucleotide primed polymerase chain reaction of purified viral nucleic acids. J Virol Methods **152**: 18-24.

Nicholas Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie WR, Hicks J and Wigler M. 2011. Tumour evolution inferred by single-cell sequencing. Nature **472**: 90–94.

Pinkney AE, Harshbarger JC, May EB and Melancon MJ. 2001. Tumor prevalence and biomarkers of exposure in brown bullheads (*Ameiurus nebulosus*) from the tidal Potomac River, USA, watershed. Environ Toxicol Chem **20**: 1196-1205.

Pinkney AE, Harshbarger JC, Karouna-Renier NK, Jenko K, Balk L, Skarphéðinsdóttir H, Liewenborg B and Rutter MA. 2012. Tumor prevalence and biomarkers of genotoxicity in brown bullhead (*Ameiurus nebulosus*) in Chesapeake Bay tributaries. *Sci Tot Environ*, *410–411:* **pp** 248-257.

Rafferty SD, Blazer VS, Pinkney AE, Grazio JL, Obert EC and LoriBoughton L. 2009. A Historical Perspective on the "Fish Tumors or Other Deformities" Beneficial Use Impairment at Great Lakes Areas of Concern. Journal of Great Lakes Research **35**: 496-506.

Smith IR, Ferguson HW and Hayes MA. 1989. Histopathology and prevalence of epidermal papillomas epidemic in brown bullhead, *Ictalurus nebulosus* (Lesuer) and white sucker, *Catostomus commersoni* (Lacépède) populations from Ontario, Canada. J Fish Dis **12**: 373-388.

van der Aa LM, Jouneau L, Laplantine E, Bouchez O, Van Kemenade L, Boudinot P. 2012. FinTRIMs, fish virus-inducible proteins with E3 ubiquitin ligase activity. Dev Comp Immunol **36**: 433-41.

Williamson SJ, Rusch DB, Yooseph S, Halpern AL, Heidelberg KB, Glass JI, Andrews-Pfannkoch C, Fadrosh D, Miller CS, Sutton G, Frazier M and Vente, JC. 2008. The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. *PLoS ONE*, **3**: e1456.

Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Khovayko O, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Ostell J, Pruitt KD, Schuler GD, Shumway M, Sequeira E, Sherry ST, Sirotkin K, Souvorov A, Starchenko G, Tatusov RL, Tatusova TA, Wagner L and Yaschenko E. 2008. Database resources of the National Center for Biotechnology Information**.** Nucleic Acids Res **36**: D13-21.

**Appendix 1 (See attached)**

**Appendix 2 (Electronic file)**

**Appendix 3 (Electronic file)**