

IDENTIFICATION OF DIFFERENTIALLY EXPRESSED PROTEINS IN
Xiphophorus FISHES OF THE GORDON-KOSSWIG
MELANOMA MODEL

THESIS

Presented to the Graduate Council
of Texas State University–San Marcos
in Partial Fulfillment
of the Requirements

for the Degree
Master of SCIENCE

by

Amy N Perez, B.S.

San Marcos, Texas
August 2006

COPYRIGHT

by

Amy Nichole Perez

2006

CHAPTER I

INTRODUCTION

The genus *Xiphophorus* consists of 26 species of freshwater, livebearing platyfish and swordtails. These species have been found in the rivers, streams, and ponds of Mexico, Guatemala, Belize, and Honduras. Interspecies hybridization and backcross interspecies hybridization between species of *Xiphophorus* has allowed researchers to examine the underlying genetic events that correspond with inheritance of particular parental phenotypes. For example, interspecies backcross hybridization between select *Xiphophorus* parental lines result in progeny (BC₁) that exhibit increased susceptibility to spontaneous or induced tumorigenesis (see below). The increased susceptibility of particular interspecies hybrids to develop tumors is likely an easily scored phenotype resulting from bringing together divergent genomes from two distinct species.

A. DNA REPAIR IN *Xiphophorus* INTERSPECIES HYBRIDS

It has been observed that first generation interspecies hybrids possess DNA repair capabilities that may resemble either one parent in

some tissues and/or exhibit reduced capabilities compared to either parents in other tissues. One hybrid model studied for base excision DNA repair capability is the F₁ hybrid produced from crossing *X. maculatus* Jp 163 A with *X. couchianus*; termed *Sd-couchianus* tumor model. *Sd* is a sex-linked genetic marker for the "spotted dorsal" pigment pattern (1,2). Base excision repair is a multi-enzymatic pathway employed to repair DNA damage caused by hydrolysis, alkylation, and attack by reactive oxidative species (3). In this cross base excision repair (BER) capabilities were assayed in tissue derived from parental and F₁ hybrid fish using oligonucleotide substrates containing a G:U mismatch (3). The *X. maculatus* parent and the *Sd-couchianus* F₁ hybrid exhibited similar repair capabilities in brain tissue but this activity was less than BER levels observed in the brain tissue from the *X. couchianus* parent. Also, the *X. couchianus* and the *Sd-couchianus* F₁ hybrid showed similar BER repair in the liver but the levels of repair were much less than was observed for *X. maculatus* liver BER. It is interesting that within this hybrid model, the *Sd-couchianus* F₁ hybrid possessed lower repair in gill tissue than was observed for either of the parental gills.

Studies using *Xiphophorus* interspecies hybrid models to examine a second DNA-repair pathway, nucleotide excision repair (NER), have also been reported (4). NER is the multi-enzymatic pathway principally employed to repair DNA damage induced by exposure to ultraviolet (UV) light and some chemicals that lead to formation of bulky DNA base adducts (4). Two types of UV-induced

DNA damage assayed were the cyclobutane pyrimidine dimer (CPD) and the pyrimidine(6-4)pyrimidone dimer [(6-4)PD]. The two *Xiphophorus* hybrid models studied were the *Sd-couchianus* and the *Sp-couchianus* F₁ hybrid, which is produced by crossing *X. maculatus* Jp 163 B with *X. couchianus*. *Sp* is also a sex linked genetic marker that is responsible for the "spotted side" pigment pattern in *X. maculatus*. It was observed that skin tissue from *Sd-couchianus* F₁ animals had reduced 6-4PD repair capabilities compared with that of either parent. Notably, *Sp-couchianus* F₁ animals exhibited repair that was reduced by an order of magnitude compared to either parent.

Changes in DNA repair capabilities among tissues in interspecies hybrids compared to the same activities in the parental lines giving rise to the hybrids suggests that species-specific allele interactions between DNA repair genes that encode the enzymes in the repair pathways may serve to modulate overall DNA repair capability in interspecies hybrid animals. These DNA repair effects are interesting considering these same *Xiphophorus* interspecies hybrid models have also been utilized to study the genetics underlying induced tumorigenesis.

B. *Xiphophorus* TUMOR MODELS

1) DNA Damage Induced *Xiphophorus* Tumor Models

Several *Xiphophorus* interspecies backcross hybrid models have been shown to develop specific types of tumors after treatment with UV light or chemical mutagens such as *N*-methylnitrosourea (MNU). One such interspecies

backcross hybrid, the *Sp-helleri* BC₁, ([*X. maculatus* Jp 163 B x *X.helleri* (Sara)] x *X. helleri*) has been shown to produce animals susceptible to UV or MNU induced melanomas. In similar fashion *Sp-couchianus* BC₁ hybrids ([*X. maculatus* Jp 163 B x *X. couchianus*] x *X. couchianus*) also show increased incidence of melanoma in pigmented classes of progeny after UV exposure 6 days post-birth or MNU exposure at 6 weeks post-birth (5, 6, 7,8,9,). In contrast to these models, the *Sp-Andersi* backcross hybrid model ([*X. andersi* x [*X. maculatus* Jp 163 B x *X. andersi*]) has been shown to have high susceptibility to MNU induced tumorigenesis (\approx 30% incidence; 10) while appearing to be refractory to tumorigenesis after exposure to UV light. These are but a few examples of *Xiphophorus* interspecies tumor models that cumulatively suggest interspecies hybridization may result in disruption of biochemical pathways related to proper cell-cycling capabilities. This may result in animals that are predisposed to tumor development and thus exposure to DNA damaging agents is capable of inducing cellular transformation. However, that different *Xiphophorus* crosses lead to varied tumor susceptibilities may also suggest that each combination of divergent alleles represented by a *Xiphophorus* interspecies model may produce different genetic effects on any particular biochemical pathway.

2) The Gordon-Kosswig Spontaneous Melanoma Model

The first *Xiphophorus* interspecies cross that was shown to produce tumor development in BC₁ hybrids consists of crossing the platyfish, *X. maculatus* Jp

163 A, carrying *Sd* with the swordtail, *X. helleri* (Sarabia), followed by backcrossing F₁ interspecies hybrid progeny to the *X. helleri* parental line (figure 1-1). This particular cross is the most widely utilized *Xiphophorus* interspecies hybrid cancer model. The cross is named the Gordon-Kosswig melanoma model (G-K model) after the two scientists who independently published its characteristics in the late 1920's (1,2). In the initial cross, F₁ hybrids have more pronounced melanin pigmentation in the dorsal fin than either of the parents. Backcrossing of the F₁ interspecies hybrid to the *X. helleri* parent results in the production of three classes of backcross hybrid progeny with respect to pigmentation. Importantly, the pigment classes of BC₁ progeny exhibit Mendelian segregation ratios expected from a two-gene model where 50% are unpigmented, 25% have a pigment pattern resembling the F₁ animals, and 25% of the progeny develop severely enhanced melanization of the dorsal fin region that will spontaneously develop into melanoma.

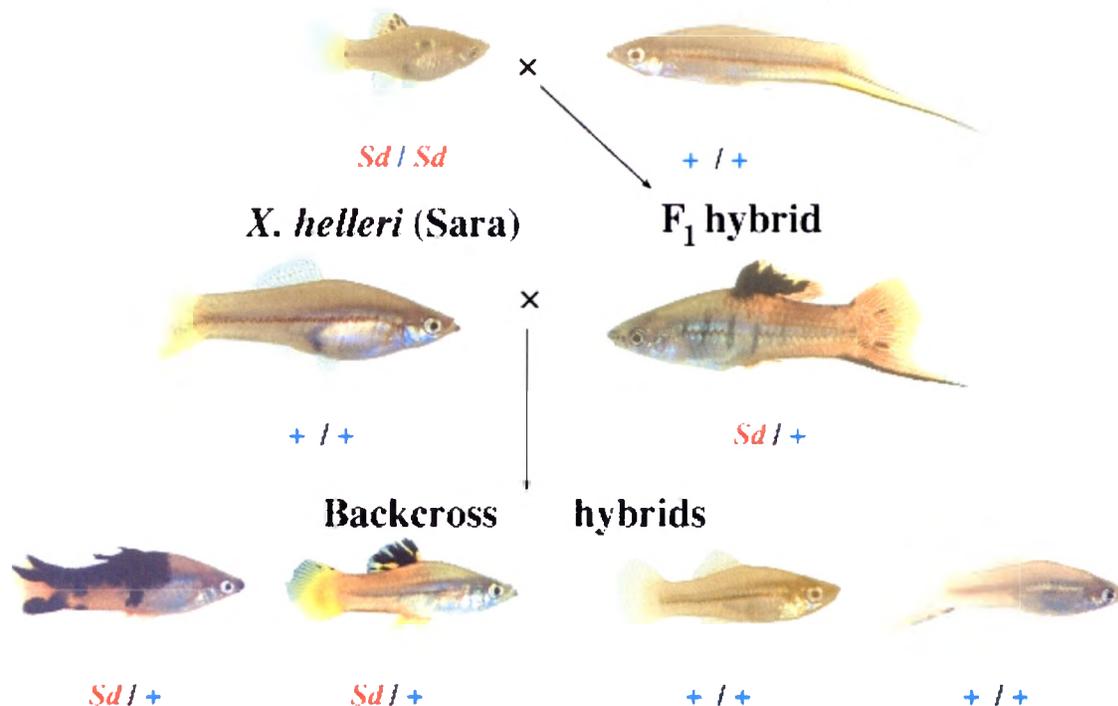
Based on a two-gene model for the G-K melanoma, one genetic marker was identified very early as a predictor of melanoma, the *Sd* locus, which is responsible for the pigmentation needed for melanoma development (1,2). Much later, a hypothetical loss of a tumor suppressor gene, *Diff*, was proposed to interact with *Sd*. The tumor factor, *Diff*, was genetically mapped to linkage group V (LGV) of *Xiphophorus* (11). More recently, Kazianis *et al.* (12) cloned a cyclin-dependent kinase inhibitor-2 (CDKN2X) and found this gene mapped very close to the proposed *Diff* tumor factor on LGV. In addition, the CDKN2X gene bears

striking homology to the human p15 and p16 genes that have been shown to be associated with human melanoma. Thus, CDKN2X has been forwarded as a candidate gene for *Diff*. However, it remains unproven if CDKN2X is the elusive *Diff* tumor factor or if it just happens to map close to it (12).

The second gene involved in the G-K melanoma model is the oncogene *Xiphophorus* melanoma receptor tyrosine kinase-2 (*Xmrk-2*). The *Xmrk-2* locus has been mapped to the X chromosome in *X. maculatus* along with its related proto-oncogene *Xmrk-1* (13,14). Both of these genes were found to be tightly linked to the *Sd* locus, the pigment pattern marker that had earlier been associated with tumorigenesis (15,16). *Xmrk-1* was shown to be expressed at low levels in all tissues in contrast to *Xmrk-2* which is overexpressed in melanomas in the tumor-bearing BC₁ hybrids of the Gordon-Kosswig cross (17). *Xmrk-2* has been shown to associate with phosphatidylinositol-3 kinase (PI3 kinase) which is known to be related to cell cycle function (18). It is still unclear if *Xmrk-2* affects only PI3 or if there are other unknown kinases effected by *Xmrk-2* that contribute to the spontaneous melanoma development. The precise genetic regulation that leads to G-K melanoma remains a very active research question.

The "Gordon-Kosswig" (H005BC₁-B) Cross

X. maculatus (Jp 163 A) *X. helleri* (Sara)



www.xiphophorus.org

Figure 1-1. A representation of *X. maculatus* Jp 163 A with the spotted dorsal (*Sd*) pigment pattern being crossed with *X. helleri* (Sarabia). The resulting progeny are interspecies *Sd-helleri* F₁ hybrids. A backcross hybridization of the *Sd-helleri* F₁ (*X. maculatus* Jp 163 A x *X. helleri*) to *X. helleri* results in three classes of backcross (BC₁) progeny with respect to pigmentation. The resulting progeny are 50% unpigmented with no *Sd* allele, 25% having pigment similar to the F₁ with one *Sd* allele, and 25% developing severe pigmentation that eventually develops into melanoma and having one *Sd* allele.

C. INTERSPECIES HYBRIDIZATION AND GLOBAL GENETIC

DYSREGULATION: A HYPOTHESIS

From the above examples of DNA repair and cell cycle dysregulation within *Xiphophorus* interspecies hybrids one may hypothesize that bringing together divergent genomes leads to global dysregulation. In this case, the initial cross producing F₁ progeny is the step in between the established parental lines and the interspecies backcross hybrids. The F₁ hybrids might be expected to show effects from crossing genetic lines on any particular trait while the backcross interspecies hybrids may amplify certain traits due to loss of non-recurrent parental alleles.

To approach this from a cancer viewpoint one can distinguish between genetic regulatory differences that are the result of the interspecies hybridization from the genetic events in target cells that result in development of cancer. If so, one may develop a cadre of valuable biomarkers for melanoma development. Past research has weighed heavily on looking at potential genes involved in carcinogenesis one at a time. This study will address interspecies hybridization and the development of melanoma from a protein (i.e. proteomic) standpoint. A proteomic approach allows us to look at an entire proteome for protein abundance differences. First, the proteome of the F₁ skin tissue will be compared to that from parental species to examine global effects on protein abundance solely due to the interspecies hybridization. Then, each proteome of backcross interspecies hybrids that have developed melanoma will be compared

to the F₁ skin proteome to attempt to identify protein differences unique to melanoma. Proteins that are continually changing expression levels with both hybridizations are likely due to the hybridization itself. Whereas, proteins that are affected only in the tumor tissue may be due to the tumorigenic process and are potential candidates to assist our understanding of cell pathway changes that could later lead to identification of melanoma biomarkers.

D. RESEARCH GOALS

The overall goal of this research is to study the G-K melanoma model using a proteomics approach. Instead of looking at one gene at a time and making a decision regarding its relevance to melanoma, a proteomic study may allow the identification of differential protein abundances that are related to two distinct biological events - interspecies hybridization and melanoma development. Direct assessment of the genetic regulatory effects resulting from allelic interplay within interspecies hybrids has not been studied previously. Utilization of the *Xiphophorus* model system for this first look at the effects of interspecies hybridization on the proteome has particular value since changes in the proteome may be followed through successive crosses. Thus, by exploring differential protein abundances due to interspecies hybridization, and then due to backcross interspecies hybridization, we will likely establish new data and genetic regulatory mechanisms that may lead to new research directions.

E. 2D- PAGE AND DIGE

One proteomic tool used in this study is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). This method involves the separation of proteins first by isoelectric focusing. Isoelectric focusing involves separation of proteins by their isoelectric point which is based on the amount of acidic and basic residues present in the protein and the overall charge based on those residues. A pH gradient is introduced and proteins migrate within that pH gradient until they stop at a pH that is equal to their isoelectric point. The second dimension involves the separation of the proteins based on their molecular weight. The two dimensional separation allows one to look at individual proteins and the quantity of each protein would be analyzed for comparison between samples based on some visualization method.

There have been problems associated with visualization of proteins on a gel and the accuracy of quantitation. One stain commonly used for visualization is the organic Coomassie Blue dye R250. The problem with R250 stain, as with other organic dyes, is that the destaining process is very hard to exactly reproduce, sensitivity of detection is low, and there is a low linear dynamic range of quantitation (19). Silver staining is another stain used that is much more sensitive than R250. However, silver stain is not quantitative due to unreliable interactions between individual proteins and the silver ions (20). Fluorescent staining has become the more preferred method. It is sensitive to low

abundance proteins like the silver stain and has a wider dynamic range than that of Coomassie Blue (20).

An issue with staining a gel is that only one sample can be analyzed on any particular gel. This introduces a problem with intra-gel variation. Intra-gel variation affects whether samples to be compared on different gels will be able to be superimposed in order for software programs to look for expression differences (21). Difference Gel Electrophoresis (DIGE) was introduced to alleviate this problem (22). The DIGE system allows different comparison samples to be run on the same gel. Fluorescent cyanine dyes, Cy3 and Cy5, which respond to different wavelengths are covalently bound to the lysines of proteins in the samples to be examined (22). These cyanine dyes are positively charged to compensate for the positive charge loss on the lysine so that the isoelectric focusing is not affected. Both cyanine dye masses are matched that introduce a slight but identical increase in mass to each labeled protein so that the separation by molecular weight is unaffected. Running comparison samples on the same gel helped to solve the intra-gel variation but inter-gel variation between several comparisons still remained.

DIGE was modified through the introduction of a third dye, Cy2, to help normalize spots across different gels. Cy2 labels equal amounts of both samples and is run along with Cy3 and Cy5-labeled samples on each gel (23). Cy2 normalizes spots across all gels, which increases quantitation accuracy and

confidence in abundance differences observed (24). The DIGE process including the Cy2 labeling is summarized in figure 1-2.

Replicate variability is a constant worry in 2D-PAGE analysis. Variation in protein levels between different extractions from the same source, multiple labelings, and gels run at different times can all contribute to the variability seen in the quantitation analyses of compared samples (25). The proteins that are reported as statistically significant in abundance in this report are shown consistently significant regardless of varying extractions, labeling timing, or replicates of gels.

Consistency seen despite all these variables helps to ensure that the biological variation seen in the proteins in the gels is due to real differences in the experimental system examined.

F. MASS SPECTROMETRY AND PROTEIN IDENTIFICATION

Once proteins have been statistically selected as differentially abundant, the proteins will be examined by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry for peptide sequencing by postsource decay (PSD) and identification through public database searching (for review of MALDI see 26).

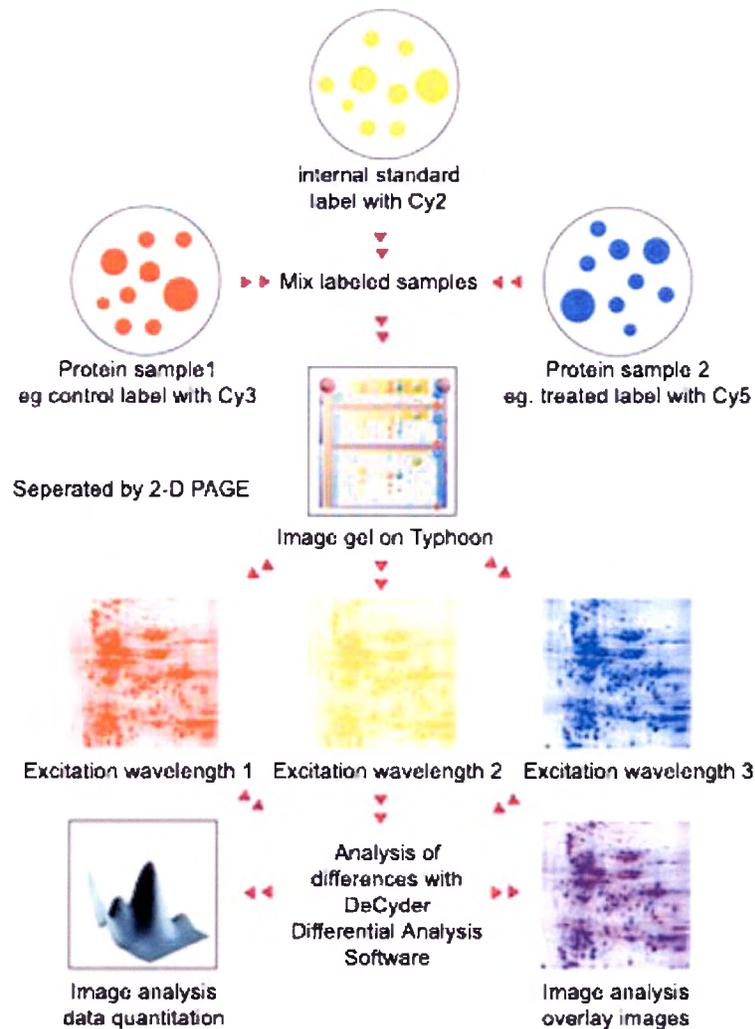


Figure 1-2 – DIGE system overview. Protein sample 1 is labeled with Cy3, protein sample 2 with Cy5, and equal amount of both protein samples are labeled with Cy2. The protein samples are separated on the same 2D-PAGE gel. The gel is imaged by a Typhoon imager and analyzed by DeCyder Analysis software.

Taken from with modifications

<http://www1.amershambiosciences.com/aptrix/upp01077.nsf/Content/Products?OpenDocument&ModuleId=165425&hometitle=search>

Problems associated with PSD for sequencing include the low signal to noise ratios of the fragments, obtaining only partial sequences, and complex fragmentation patterns (27). Chemical modifications to peptides are one way of improving upon these issues. We have chosen N-terminal derivatization of our peptides by 4-sulfophenyl isothiocyanate (SPITC) (28). SPITC was shown to improve the PSD spectra by simplifying the fragmentation of peptides to give a predominately y-ion series providing one major fragmentation pattern (27,28). Y-ion fragments produced from this derivatization in combination with fragmentation have high signal to noise ratios and generally give complete sequences (27,28). The mass differences of the peptide fragments match the masses of amino acids that make up the particular peptide. The amino acid sequence for each peptide is searched against public protein databases for tentative identification.

H. SUMMARY

Here is presented the validation of the DIGE method for examining the proteomic responses due to *Xiphophorus* interspecies hybridization. Proteins obtained from dorsal fins of the fishes involved in the Gordon-Kosswig model are used to identify proteins differentially abundant due solely to hybridization and due specifically to melanoma development.

CHAPTER II

MATERIALS AND METHODS

A. MATERIALS

All bulk chemicals that are not explicitly listed here were analytical-grade or better and were purchased from Invitrogen (Grand Island, NY). The cyanine fluorescent dyes (CyDyes), pH 3-7 non-linear immobilized pH gradient (IPG) strips, isoelectric focusing (IEF) rehydration buffer, Bind-Silane, and Deep Purple stain were purchased from Amersham (GE Healthcare, Piscataway, NJ). HPLC-grade acetonitrile was purchased from Burdick-Jackson, Morristown, NJ. Sequencing-grade trifluoroacetic acid (TFA), *N,N,N',N'*-tetramethylethylenediamine (TEMED), ProteoPrep Chaotropic Extraction Reagent 3, 4-sulfophenyl isothiocyanate, 2,4,6-trihydroxyacetophenone monohydrate (THAP), and α -cyanohydroxycinnamic acid (HCCA) were purchased from Sigma (St. Louis, MO, USA). ZipTip reversed-phase pipet tips were purchased from Millipore (Bedford, MA, USA). The peptide standards used to calibrate the mass spectrometer were purchased from Bruker Daltonics (Billerica, MA).

B. SAMPLE PREPARATION

Dorsal fin samples from 25 parentals *X. maculatus* Jp 163 A (generation 102), 30 *Sd-helleri* F₁ hybrids, and 16 *Sd-helleri* BC₁ hybrid tumors were provided by the *Xiphophorus* Genetic Stock Center, Texas State University, San Marcos, TX, USA. Each fish was placed in 0.1% MS-22 until their activity slowed. The fish were then laid on a glass plate and their dorsal fins removed with a scalpel. The fish without dorsal fin were returned to their aquaria. Dorsal fins were taken from the same fish on two other occasions. These fish are still alive in the event they are needed for further experimental use. Dorsal fins of each experimental group were pooled and total proteins extracted using 200 μ l Sigma ProteoPrep Chaotropic Extraction Reagent 3. After homogenization using a handheld pestle, the proteins were sonicated five times for 15-20 seconds, cooling on ice between sonications. The protein samples were centrifuged at 13,000 x g for 30 min. at room temperature (rt). The supernatant was transferred to a fresh bullet tube and reduced with 200 mM tributylphosphine at a final concentration of 5 mM. The samples were incubated at rt for 1 hr, then alkylated with 0.05 M iodoacetamide at a final concentration of 15 mM. The protein samples were incubated at rt for 1.5 hrs in the dark. After centrifugation at 13,000 x g for 15 min., the protein supernatant was transferred to a new centrifuge tube.

Sample protein concentrations were determined using the Bio-Rad Bradford protein assay (Hercules, CA). The bovine serum albumin (BSA) standards were prepared in water with a working range of 200 – 1.56 μ g/ μ l. Each standard was prepared by adding 200 μ g/ μ l in the first well of a microplate

and then serially diluting until reaching $1.56 \mu\text{g}/\mu\text{l}$ with a total of $100 \mu\text{l}$ in each well. Each sample well was prepared by adding $10 \mu\text{l}$ of sample to $190 \mu\text{l}$ of deionized water and then serially diluted until a final dilution contained a ratio of 1:640 sample:water. There was also a blank well containing only water. At this point each of the wells contained $100 \mu\text{l}$. The dye reagent concentrate provided in the Bio-Rad Bradford kit was diluted 2:3 with water. Dye reagent was added in $100 \mu\text{l}$ aliquots to all standard and sample wells. Each well was mixed by pipetting and the plate incubated at rt for 5 min., followed by assessment of absorbance in each well at 595 nm using a BIO-TEK Powerwave microplate scanning spectrophotometer. The reading from the water blank was subtracted from all measurements. A standard curve was prepared by plotting the adjusted absorbance measurement to known concentrations of BSA protein. Protein concentrations for each sample were determined based on the standard curve.

C. DYE STOCKS AND WORKING SOLUTION

The CyDyes (Cy3, Cy5, and Cy2) were reconstituted in dimethyl formamide (DMF) to a final amount of 1 mM. A $200 \text{ pmol}/\mu\text{l}$ CyDye working solution was prepared by adding $4 \mu\text{l}$ of DMF to a fresh microfuge tube followed by $1 \mu\text{l}$ CyDye stock solution. This was done for each of the three dyes.

D. ISOELECTRIC FOCUSING SEPARATION

Once the fin protein samples were prepared, DeStreak rehydration solution used in the rehydration of the IEF DryStrip gels was equilibrated to rt and $15 \mu\text{l}$ of IPG buffer added per 3 ml of rehydration solution. pH 3-7 non-linear

Immobiline DryStrips were rehydrated in a reswelling tray with DeStreak rehydration solution containing IPG buffer and labeled protein. The labeled protein mixture contained (1) *X. maculatus* Cy3-labeled sample, (2) *Sd-helleri* F₁ Cy5-labeled sample, and (3) Cy2-labeled standard sample. The Cy2-labeled sample contained equal amounts of both protein samples to be compared in the subsequent DIGE gels. The Cy2-labeled sample is used for normalization across multiple gels. This dye labeling scheme was repeated once. Two additional protein mixtures were prepared identically except they contained “dye flipped” samples where *X. maculatus* was labeled with Cy5 and *Sd-helleri* was labeled with Cy3. An equal amount of rehydration solution was added to each of the four protein mixtures. Each protein mixture containing added rehydration solution was used along with additional rehydration solution to rehydrate the IEF strip. The gels were allowed to rehydrate overnight at rt.

After rehydration, the IEF strip containing the protein samples was rinsed with water and loaded on Amersham Ettan IPGphor II IEF unit for first dimension isoelectric focusing. The manifold tray was filled with 108 ml of mineral oil. The strips were transferred face up into the tray with the anodic end of the IPG strip resting at the top of the tray. Each strip was centered down the entire length of the channel. Individual paper wicks hydrated with 150 μ l of water were placed at the ends of each of the strips. Electrodes were placed on the wicks and the electrode cams closed. A cover was put on the machine to exclude light. The recommended running conditions were as follows:

Step	Action	Volts	mA	Volt hours	Hours
A	Step and Hold	300V	50 μ A	900 Vh	3 hrs
B	Gradient	600V	50 μ A	1350 Vh	3 hrs
C	Gradient	1000V	50 μ A	2400 Vh	3 hrs
D	Gradient	8000V	50 μ A	13500 Vh	3 hrs
E	Step and Hold	8000V	50 μ A	3200 Vh	4 hrs
F	Step and Hold	500V	50 μ A		

E. 2D-PAGE GEL PREPARATION

During isoelectric focusing, 12.5% SDS-PAGE gels were cast. These large format gels are 25.5 x 20.5 cm and 1 mm thick. Prior to casting the gels, each gel plate is soaked overnight in 5% Decon 90 solution. The next day the gel plates are rinsed of the soap and cleaned with 3% acetic acid followed by 100% ethanol. If the gels are for preparative purposes, the glass plates (without spacers) are treated with Bind Silane (GE Healthcare, Piscataway, NJ) so the gels will bind to the plates during the robotic spot-picking process. Bind-Silane working solution consisted of 80% ethanol, 1.9% glacial acetic acid, 0.1% Bind-Silane, and 18 μ l of water. Once the plates were dry after cleaning, 2 mL of Bind-Silane solution was pipetted onto the surface of each plate and spread evenly to cover the entire plate. Each treated plate is allowed to dry for one hour. After that hour, two self-adhesive fluorescently detectible reference markers are attached to the treated plates prior to gel casting. Before pouring the acylamide gel solution, the gel caster is assembled with four sets of low-fluorescence glass plates and two plastic spacers that take up the same amount of space as glass

plates. For gel caster assembly refer to the user manual Ettan Daltsix Electrophoresis System by GE Healthcare.

For each gel, 41.7 ml monomer stock solution (30% acrylamide, 0.8% N,N'methylenebisacrylamide), 25 ml 4X resolving buffer (1.5 M Tris-Cl, pH 8.8), 1 ml 10% SDS, and 31.8 ml double distilled water is used. This gel suspension is filtered through a 0.45 μm filter, then 500 μl of 10% ammonium persulfate and 33 μl of TEMED were added to the gel suspension and quickly mixed. The gels are poured immediately upon mixing. Once the solution is poured into the caster, leaving a 1 cm gap from the top of the shorter glass plate, the tops of the cassettes are sprayed with 0.1% (w/v) SDS/water using a hand-held spray bottle. These gels are cast at least one day before running to ensure complete polymerization. Each gel is removed from the caster, rinsed with distilled water (to remove the SDS), and placed on a holding rack until use.

F. SDS-PAGE

After isoelectric focusing each IEF strip is equilibrated in 2 ml of equilibration buffer (50 mM Tris-Cl, pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, plus a few grains of bromophenol blue) containing 2 mM DTT for 10 min. Then the IEF strips are transferred to equilibration buffer with 2 mM iodoacetamide and equilibrated for 10 additional min.

During the equilibration, SDS electrophoresis running buffer is prepared (25 mM Tris base, 192 mM glycine, 0.1% SDS, in distilled water). Once the IEF strips are equilibrated, they are rinsed in 1X SDS running buffer and immediately layered on the top of 12.5% SDS-PAGE gels. To ensure uniformity of the PAGE

gels, the IEF strips are placed on the PAGE gel with the strip gel facing forward and the acidic end on the left side when the shorter plate is facing forward. Once the strips are in contact with the PAGE surface, each strip is sealed to the PAGE gel with 0.5% agarose overlay solution. The lower buffer tank of the Amersham DALTSix electrophoresis unit is filled with 4 L 1X SDS electrophoresis buffer and the gel rack inserted to the tank. The upper buffer chamber is seated over the gel rack and filled with 2X SDS electrophoresis buffer. The gels are electrophoresised at 5 (watts/gel) for 30 min. followed by 17 (watts/gel) for 4 hrs (until the dye front touches the bottom of the gel).

G. GEL IMAGING

Following the second dimensional separation each gel plate is rinsed with water and then with ethanol to remove SDS residue. During plate cleaning, the Amersham Typhoon Trio apparatus is allowed to warm up for 30 min. The glass of the Amersham Typhoon Trio imager is cleaned with 10% hydrogen peroxide, followed by water, and then ethanol. Two gels can be imaged at the same time using an Ettan DALT gel alignment guide (for review, refer to Typhoon User's Guide v3.0, 2002). The alignment guide places the glass plate sandwiched gels 0.2 mm above the glass to avoid optical interferences.

Once the gels are in place, the scanner control window is opened using Typhoon Scanner Control (v5.0). The acquisition mode is set at fluorescence. The setup button in the acquisition mode box is pressed to set fluorescence parameters. Three emission filters are chosen, one corresponding to each of the three dyes (Cy2, Cy3, and Cy5) and the OK button is pressed. In the options

window, the gel orientation is chosen based upon the orientation of the gels on the imager surface. The “press sample” box is checked so that the sandwich gels are held by an inner lid to ensure no movement of the gels that may distort imaging. The initial scan of the gels is performed at 1000 μm pixel resolution. The focal plane is changed to +3 mm. The DIGE file naming format box is checked and the scan button is clicked. The program then prompts for one to add the name of the gels to be scanned. The appropriate standard is also chosen at this point (Cy2). Once the files for both gels are named, the OK button is pressed and scanning of each gel with all three lasers is performed. Using ImageQuant software the images are assessed to make sure the gels ran accurately. If the gel images exhibit expected protein spot patterns and minimal streaking, the acquisition mode setup menu is modified by changing the photomultiplier tube voltage to 580 for each of the three lasers in order to avoid detector saturation. In addition, the pixel resolution is set to 100 μm to achieve the higher resolution needed for the DeCyder image analysis software. The DIGE file naming format box remains checked and when the scan button is pressed files for each gel are appropriately named as stated.

H. DECYDER ANALYSIS

Once all gels are scanned, each gel image is processed for DeCyder analysis. A gel image of a combination of all three dye images for each gel scanned is brought up on the computer screen using ImageQuant TL software. The “edit image” button is selected and the image is flipped horizontally and saved. The flipped saved image is then opened in ImageQuant 5.1. Under the

tools menu “define region of interest” is selected. This tool is used to outline the gel image while excluding any part of the alignment guide that was imaged. The outlined image is saved and used within the DeCyder software for image analysis.

When the DeCyder analysis program is opened, Image Loader is the first icon chosen. For each gel run, a new project is created within Image Loader. Gel images are added to the import list by clicking the “add” button and selecting all three gel images of the same gel and then clicking “open”. For each gel, the edit button is selected to make sure the dye chemistry is DIGEmin (minimal CyDye labeling) and each of the three fluorescent images of that gel is correctly labeled according to its dye assignment. Once those settings are checked and the “Ok” button is pressed, the images may be put into the appropriate project folder by clicking the import button. Image Loader is then closed and the Differential In-gel Analysis (DIA) module icon is selected.

In DIA, the Create Workspace button is pressed and all three images of the first of the gels is selected. The “create” button is clicked to create the DIA workspace. After the images have been loaded, the process menu is selected followed by selecting Process Gel Images. In the Process Gel Images dialog box, 2500 is selected as an estimated number of spots. If the gels are preparative, the Autodetect Picking references is checked. The “Ok” button is then selected to begin spot detection. After the computer selects spots, the images are visually inspected to ensure that all spots were detected. Reference locations are edited if there are any discrepancies between where the software

located the references and where the references are actually located on the image. The rest of the gels are each processed individually like the first.

Once all the gels have been processed in DIA, the DIA program is closed and the Biological Variation Analysis (BVA) module icon selected. Once the module is opened, a workspace is created and two DIA workspaces are added corresponding to two replicate gels of the same dye chemistry from one DALTSix gel run. Once these workspaces are added, the Spot Map Table mode is automatically brought up. In the Experimental Design View panel, the “Add” button is clicked and a folder is named after one of the samples and then confirmed. This is repeated to add another folder corresponding to another sample. All gels files are originally in the control folder. After the other two folders are created, the gel files corresponding to each sample are moved to the appropriate folder. Next, the “Match Table” mode is selected and Cy2 images for both gels to be compared are selected. The Landmark button is clicked and spots are selected in the “master image” and then selected in the “secondary image” for matching. Twenty spots are selected by the user that are common between all gels and cover all regions of the gels. Once 20 spots are matched the Landmark button is deselected.

In the Process menu Match is selected. In the Match window, the “Match All” option is selected and the match button clicked. After the images are matched the “Match Table” button will need to be pressed once more. The match vectors are checked to ensure accurate matching by looking to see that the vectors are all facing the same direction. About ten Auto Level 1 matches,

the best matches assigned by the software, were checked for accuracy as well as about 20 Auto Level 2 matches, the second-best matches assigned by the software. The matches are checked visually to see if the computer accurately selected corresponding protein spots in both gels. If the protein spots were matched correctly between gels, the Protein Table (PT) icon is selected. The Process menu is selected and "Protein Statistics" chosen. For example, the parent samples are chosen for population 1 and the F₁ hybrid offspring selected as population 2. The boxes checked included the independent tests (normal), average ratio, Student's T-test, and One-Way ANOVA between different groups. The "calculate" button is clicked and the Process menu selected with the Protein Filter Dialog chosen. The Assign Proteins of Interest box is selected. Then the Student's T-test p-value is set at 0.05 and the average ratio set at <-1.5 to >1.5. Once all the boxes are checked, the filter button is clicked followed by the OK button. This protein filter provides proteins of interest (POIs) that pass all statistical criteria. For easier identification of POIs, the view menu is selected and the properties dialog box chosen. The protein table tab is clicked and in the Protein Table Filter, Proteins of Interest are selected so that only proteins identified in the protein table and on the gel images are the POIs. This entire analysis process is repeated with both two dye flip gels harboring proteins differentially labeled from the same sources.

The entire DIGE analysis using the same two samples is repeated two more times. In each analysis there are POIs for one dye labeling scheme and a set of POIs for the dye flip. At the end of all the runs, there are six sets of BVA

analyses consisting of three experiments with two duplicate gels of two dye labeling schemes each. Each of the spot maps with POIs outlined is printed out for visual comparison. The spot maps of one dye scheme are analyzed to find POIs that are found to be consistently up or down-regulated. This visual analysis is done for the dye flip spot maps as well. Once a composite of all proteins found to be consistently up or down-regulated for all three runs for each dye labeling scheme is compiled, both composites are compared for similarity. The POIs that are found in all three experiments across dye-flips are selected for further analysis and identification.

I. PREPARATIVE GELS AND SPOT PICKING

For identification by MALDI-TOF MS, preparative gels with higher protein loads than analytical gels are run in order to excise the POI spots out of the gels. Each of four preparative gels has a protein load of 400 μg , half of which is the experimental sample and half of which is the control sample. The preparative gels are run in the first and second dimensions exactly as the analytical gels. After the second dimension, the glass plate to which the gel is not bound is removed and the glass plate with the gel attached is submerged into 500 ml fixing solution of 7.5% (v/v) acetic acid / 10% (v/v) methanol. The preparative gels are left in the fixing solution overnight at rt. The following day, the fixation solution is poured off and replaced by 500 ml wash solution that contains 35 mM sodium hydrogen carbonate and 300 mM sodium carbonate. The gel is left in the wash solution for 30 min. with gentle agitation and then the wash solution is poured off and replaced with 500 ml of water. To the water, 2.5 ml of Deep

Purple stain is added. The gel is allowed to incubate in the stain solution for 1 hr. in the dark with gentle agitation. The stain solution is poured off, replaced with destain solution (500 ml of 7.5% acetic acid), and allowed to destain for 15 min. with gentle agitation. The destain solution is poured off and replaced with fresh destain solution. After 15 min. of gentle agitation, each gel is ready to be scanned. Each gel is scanned by the Typhoon Trio with an emission filter corresponding to the Deep Purple Stain. The image is scanned at a pixel resolution of 100 μm or the maximum resolution needed for analysis by the DeCyder software. The image is manipulated in ImageQuant TL and ImageQuant 5.1 as detailed previously for the analytical gels.

In the DeCyder software, each preparative gel image is loaded in Image Loader. The dye chemistry is post stain. Each preparative gel image is added to the project containing the BVA to be used for spot picking. The preparative gels are analyzed in DIA in order to detect spots and locate picking references. Then the BVA workspace used for spot picking is opened. The DIA workspaces of the preparative gels are added to the BVA workspace. The preparative gels are deselected as analytical spot maps and selected as pick spot maps in the Spot Map Table mode. The preparative gels are landmarked in the Match Table Mode as stated earlier in the protocol. To perform the matching, instead of checking the "Match All" selection, the "Match Pending and Landmarked" selection is checked and the match button clicked. The match vectors are checked as before. Once the match process is completed, the Protein Table mode is chosen. The master analytical gel is brought up and the proteins selected from

the earlier process are marked as “pick”. Any pick proteins that were picked in the master gel and that matched in the preparative gel are located in the preparative gel. Under the File menu, the Export Pick List is chosen. The pick list for each preparative gel is exported as a text file containing coordinates for picking for use by the Ettan Spot Picker. The BVA is then saved and closed.

To set up the Ettan Spot Picker, 96-well microplates are placed in the plate tray. The gel to be picked is submerged in 500 ml of deionized water in the Spot Picker holding tray. One preparative gel is placed face up into the tray with the reference markers lining up within the parallel lines in the center of the tray. The gel holders are then tightened on the edges of the gel to ensure there is no movement of the gel during the picking process. The computer Ettan Spot Picker software is then opened. Under the tools menu, “Prime Syringe” is selected and the syringe is primed with water 15 times. Then in the System menu, “system setup” is selected. The Z-position of the picker head has to be adjusted for each preparative gel so that the POIs can be excised without scratching the glass. The selections are then saved and exited. In the main window, “Load Pick List” is clicked. Once the pick list for the preparative gel is chosen, the “open button” is chosen. The pick list and coordinates of spots to be picked are loaded and the “next” button is clicked in the Load Pick List window. The next step is locating the reference markers. The “Auto-detect” button is selected and the picker automatically finds the location of the reference markers. The “next” button is pressed again and then the software indicates how much liquid will be required for the picking. The file location for the result of the picking is chosen and the file

named along with naming of the microplates. The picking begins. Once the picking is finished, the microplates are removed and the pick gel is removed. The picking process is repeated for each of the remaining preparative gels.

J. TRYPSIN DIGESTION AND PEPTIDE SULFONATION

Once all the spots have been picked, the water is removed from each of the microplate wells and the spots from all the gels are combined according to their spot number in 0.5 ml centrifuge tubes. The spots are destained in 200 μ l of 50 % acetonitrile (ACN)/ 100 mM ammonium bicarbonate (ABC) overnight. The following day the solution is removed and the gel plugs are rinsed two times with 200 μ l of 50% ACN/ 100 mM ABC. The solution is removed and the gel plugs are dehydrated for 5 min. in 100% ACN at rt. The gel plugs are dried in a vacuum centrifuge for 15 min. to remove the ACN. While the gel plugs are drying, Trypsin Gold is resuspended to 1 μ g/ μ l in 50mM acetic acid and then diluted in 40mM ABC/10% ACN to 20 μ g/ml. The dried gel plugs are resuspended in 25 μ l of the trypsin solution at rt for 1 hr. Once the gels are rehydrated, 40 mM ABC/10% ACN is added to cover the gel plugs. The samples are incubated at 37°C overnight. The next day, 100 μ l of HPLC-grade water is added each tube for 10 min. with frequent vortexing. The supernatant is removed and saved in a fresh tube labeled according to sample number. The gel slice digests are extracted twice with 50 μ l of 50% ACN/ 0.5% trifluoroacetic acid (TFA). The digests are incubated in this mixture for 1 hr each at rt after which the supernatant is removed to its corresponding sample centrifuge tube. Once all

the extracts are pooled, the samples are dried in a vacuum centrifuge for about 3 hrs. The samples are then stored at -80°C .

The peptides are dissolved in $2\ \mu\text{l}$ of $10\ \text{mg/ml}$ 4-sulfophenyl isothiocyanate (SPITC) in $20\ \text{mM}$ sodium carbonate pH 9.5 that included $4\ \text{mM}$ n-octyl- β -D-glucopyranoside (GDP). The pH of the peptides are checked and, if necessary, modified in $0.2\ \mu\text{l}$ increments with $100\ \text{mM}$ sodium carbonate (pH 9.5) until the pH of the solution falls between 8 and 9. The samples are allowed to SPITC label for 1 hr at rt. ZipTips are equilibrated with 50% ACN/0.1% TFA, 0.1% TFA, and then water. After SPITC modification, the peptides are bound to the ZipTips by repeated aspiration (25 times). The peptides are desalted with five washes of water followed by five washes with 0.1% TFA. The peptides are eluted into centrifuge tubes with $5\ \mu\text{l}$ of 50% ACN/ 0.1% TFA and then dried in a vacuum centrifuge. The peptides are redissolved in $1\ \mu\text{l}$ of 10% ACN/ 0.1% TFA.

K. MALDI-TOF/TOF-MS ANALYSIS

Two major matrix stock solutions are prepared including $20\ \text{mg/ml}$ 2,4,6-trihydroxyacetophenone (THAP) and $10\ \text{mg/ml}$ diammonium citrate (DAC) in 50% ACN/50% water. The other matrix stock solution prepared are $10\ \text{mg/ml}$ α -cyano-4-hydroxycinnamic acid (CHCA) in 50% ACN/ 0.1% TFA. These two stock solutions are combined in a ratio of 2:1 THAP/DAC:CHCA. The first step for each analysis is the addition of $0.5\ \mu\text{l}$ of the final matrix solution to a matt steel target plate spot and it is dried under a steady current of air. To each spot $0.5\ \mu\text{l}$ of peptide sample is added on top of the dried matrix. Another $0.5\ \mu\text{l}$ of matrix is

then added to the top of the peptide sample. Each spot is dried with a current of air. Once the spot is dry 1 μ l of 0.1% TFA is added and then dried. The spot is washed with 1 μ l of 0.1% TFA by pipetting up and down five times with eventual removal of the 1 μ l and drying of the spot. This washing and drying with 0.1% TFA is repeated once. The spot is now ready for MALDI analysis.

All MALDI mass spectra are obtained with a Bruker Daltonics Autoflex II TOF/TOF mass spectrometer with a 337 nm nitrogen laser. Peptide ion spectra for each sample spot are collected at a frequency of a 50 Hz repetition rate. Each peptide ion spectra is analyzed and precursor ions selected for further postsource decay (PSD) fragmentation. Each precursor ion is selected with a timed ion gate with a resolution of 100. PSD spectra of 2000 laser shots is collected for each precursor ion chosen. The PSD spectra is used to sequence each peptide fragment represented by the precursor ion. The SPITC labeling allows an abundance of y-ion fragments to be detected. The mass differences between each of these y-ion fragments equals the mass of an amino acid. Each peptide spectra is sequenced by working across the spectra and determining the amino acids that make up the sequence of y-ion mass differences.

Each amino acid sequence obtained was (1) BLAST searched for short-nearly exact matches using the NCBI non-redundant databases of *Vertebrata* and *Danio rerio*, (2) searched against the Swiss-Prot database, and (3) searched using FASTA from the University of Virginia. The best (i.e. lowest) "expect score" is accepted as the potential identification for each amino acid sequence.

CHAPTER III

VALIDATION OF DIFFERENCE GEL ELECTROPHORESIS

A. INTRODUCTION

One of the goals of this study was to determine methods to validate the statistical significance of different abundances observed for protein spots that would be complimentary to the DeCyder software analyses. The proteins of interest (POIs) considered significant had to be further validated since the DeCyder program only considers Type I errors (i.e. Type I error is when a false positive occurs and a significant change is detected when none is actually present [29]) through an alpha score. A t-test requirement was set with $\alpha \leq 0.05$ for a POI to be considered significant with 95% confidence. This alpha requirement accommodates for type I errors.

Methods we developed accommodate the concern of Type II errors that are not addressed by the DeCyder software. Type II error is when an analysis fails to detect a change when one is present (29). Type II error is addressed through the use of statistical power which is $1 - \beta$ where β is Type II error. The validation results for our different analyses are designed to maintain an alpha of $\alpha \leq 0.05$ but add an additional requirement of 80% power. Having an 80% power is a general requirement for routine statistical analyses (30). Power depends on

several factors including variance, alpha significance level, the abundance difference to be analyzed, and the sample size (29).

The variance may be kept low with good technical reproducibility but even so variance cannot be eliminated (31). The alpha level of ≤ 0.05 is a standard statistical cutoff that if increased helps to decrease Type II errors but concurrently increases the probability of Type I errors. Because of this the number of experimental replicates tends to be the easiest variable one may adjust to develop robust statistical validation (29).

Analyses were initially conducted with the DeCyder software to obtain numbers of POIs to give us a range of available proteins adjusted for the added Type II error requirement. This validation study examines how many replicates one may need to identify proteins with abundance ratios that are significant differences of ± 1.5 fold or larger. This study also examines what abundance ratios are considered significant should the number of replicates be held constant due to time and monetary concerns.

This chapter presents numbers of POIs based on their abundance ratios, spot data of matching abilities to examine technical variability, and development of a way to validate the statistical significance of protein abundance differences that addresses Type I error and Type II errors.

B. METHODS

For the first comparison, a 50 μg protein sample from *X. maculatus* and from *Sd-helleri* was labeled with 200 pmol of cyanine-3-fluorescent dye (Cy3)

and Cy5, respectively. This dye labeling scheme was repeated once. The Cy3 and Cy5 labeling assignments were reversed in two subsequent labelings as "dye flips". A control sample combining 25 µg of protein from *X. maculatus* and 25 µg of protein from *Sd-helleri* was labeled with Cy2 for use as an internal standard for each gel. In the second experiment, the dye labelings were conducted the same way except the protein samples were of *Sd-helleri* and BC₁ tumor. The reactions were incubated on ice for 30 min. in the dark, quenched with 1 µl of 10mM lysine, and incubated on ice for 10 min. in the dark.

The dorsal fin proteome comparison of *X. maculatus* versus *Sd-helleri* was conducted in 12 replicate gels. There were three separate analyses (termed A, B, and C) conducted using the DeCyder program : A - 6 gels of one dye labeling scheme (Cy3-labelled *Sd-helleri* and Cy5-labelled *X. maculatus*), B - 6 gels of the opposite dye labeling scheme (Cy3-labelled *X. maculatus* and Cy5-labelled *Sd-helleri*), C – all 12 gels including the six gels used in analysis A and the six gels used in analysis B.

Significant protein spots in each of the three analyses had to meet the following criteria of the DeCyder software: (1) protein spots had to be found in all gels of an analysis; (2) protein spots had to exhibit abundance ratios starting at 1.5 fold or greater between differentially labeled samples; (3) protein spots had to have a t-test score of $\alpha \leq 0.05$. If a protein spot in these analyses met these three criteria they were termed a protein of interest (POI).

The dorsal fin proteome comparison of *Sd-helleri* versus BC₁ hybrid tumor tissue was also conducted in three separate analyses (termed D, E, and F). The

three separate analyses included: D - 6 gels of one dye labeling scheme (Cy3-labelled *Sd-helleri* and Cy5-labelled BC₁ tumor), E - 6 gels of the opposite dye scheme (Cy3-labelled BC₁ hybrid and Cy5-labeled *Sd-helleri*), F – 12 gels including the six gels used in analysis D and the six gels used in analysis E.

Validation of statistically significant POIs detected by the DeCyder software involved the key variables in Eq. (1) (32).

$$N = 1 + 2C(CV/D) \quad (1)$$

where N is sample size, C is constant value (determined by α and β), CV is coefficient of variation, and D is the natural logarithm of observed ratio between two groups. (32 with some modifications). With $1-\beta = 0.8$ (80% power, type 2 error) and $\alpha \leq 0.05$ (type 1 error), the constant value is 7.85 (30).

A coefficient of variation (CV) was determined for individual protein spots that were matched in all gels by first calculating an average spot volume from individual normalized spot volumes that were calculated in each gel in an analysis. A standard deviation was determined for each protein spot. The standard deviation was divided by the average normalized volume of each protein spot to determine a coefficient of variation for that individual protein spot. Those individual coefficients of variation for all the protein spots were pooled to obtain an average coefficient of variation that was used in equation 1. One variable examined was the number of gels required to identify protein spots that exhibited significant abundance ratio of +/- 1.5 fold. Another variable examined

was the abundance ratio required to maintain statistical significance based on a fixed experimental replicate number.

C. RESULTS

1) Abundance Ratio Versus Number of Proteins of Interest

Figure 3-1 illustrates that the decrease in number of POIs occurred upon increasing average abundance ratio requirements in the analyses (A, B, and C) of *X. maculatus* versus *Sd-helleri* F₁.

In figure 3-1, panel A, 35 proteins qualify with at least +/- 1.5 fold difference in abundance which is about half the number of proteins (68, panel B) that qualify in the opposite dye scheme. The analysis, shown in figure 3-1, panel B, identified one protein spot with an abundance ratio of 5.25 where the largest abundance ratio detected in the dye flip (figure 3-1, panel A) was 3.50. There were 23 spots that qualified when all 12 gels were analyzed (figure 3-1, panel C) with the largest significant abundance ratio represented of 3.75. In all three comparisons (panels A, B, and C), increasing the abundance ratio cutoff served to drastically decrease the number of POIs. However, this leveled off to a few (3,7,1) proteins for ratios of 3 or above.

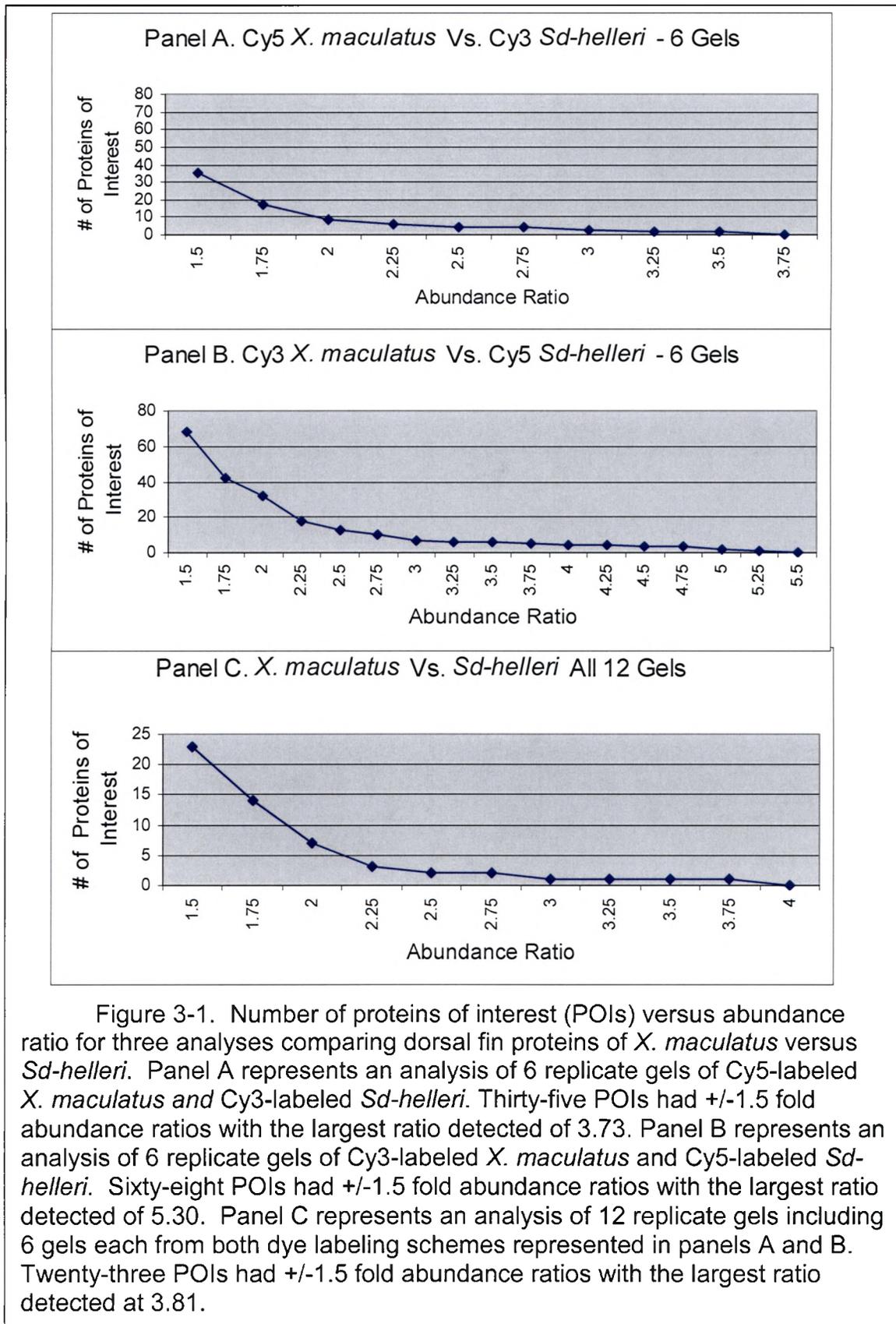


Figure 3-2 illustrates the decrease in the number of POIs one may observe due to increasing the abundance ratio requirements in the analyses of *Sd-helleri* fin vs. BC₁ tumor fin samples (D, E, and F). In figure 3-2, panel D, 113 proteins qualified as significant at an abundance ratio of +/- 1.5 whereas 91 proteins qualified in the analysis of the opposite dye scheme (figure 3-2, panel E). Figure 3-2, panel D identified two proteins with large abundance ratios (6.18,8.24). The largest abundance ratio, 5.05, was detected in the opposite dye scheme analysis (figure 3-2, panel E). There were 50 spots that qualified when all 12 gels were analyzed with the largest abundance ratio identified at + 7.30 in BC₁ tumor dorsal fin tissue compared to *Sd-helleri* dorsal fin. The three analyses (D, E, and F) indicate that the number of POIs decrease dramatically as the abundance ratio cutoffs are increased but tend to level off at abundance ratios of 3.00. Figure 3-2, panels D and E show 9 proteins of interest that exhibit abundance ratios of at least 3.00 whereas panel F (all 12 gels) shows only 4.

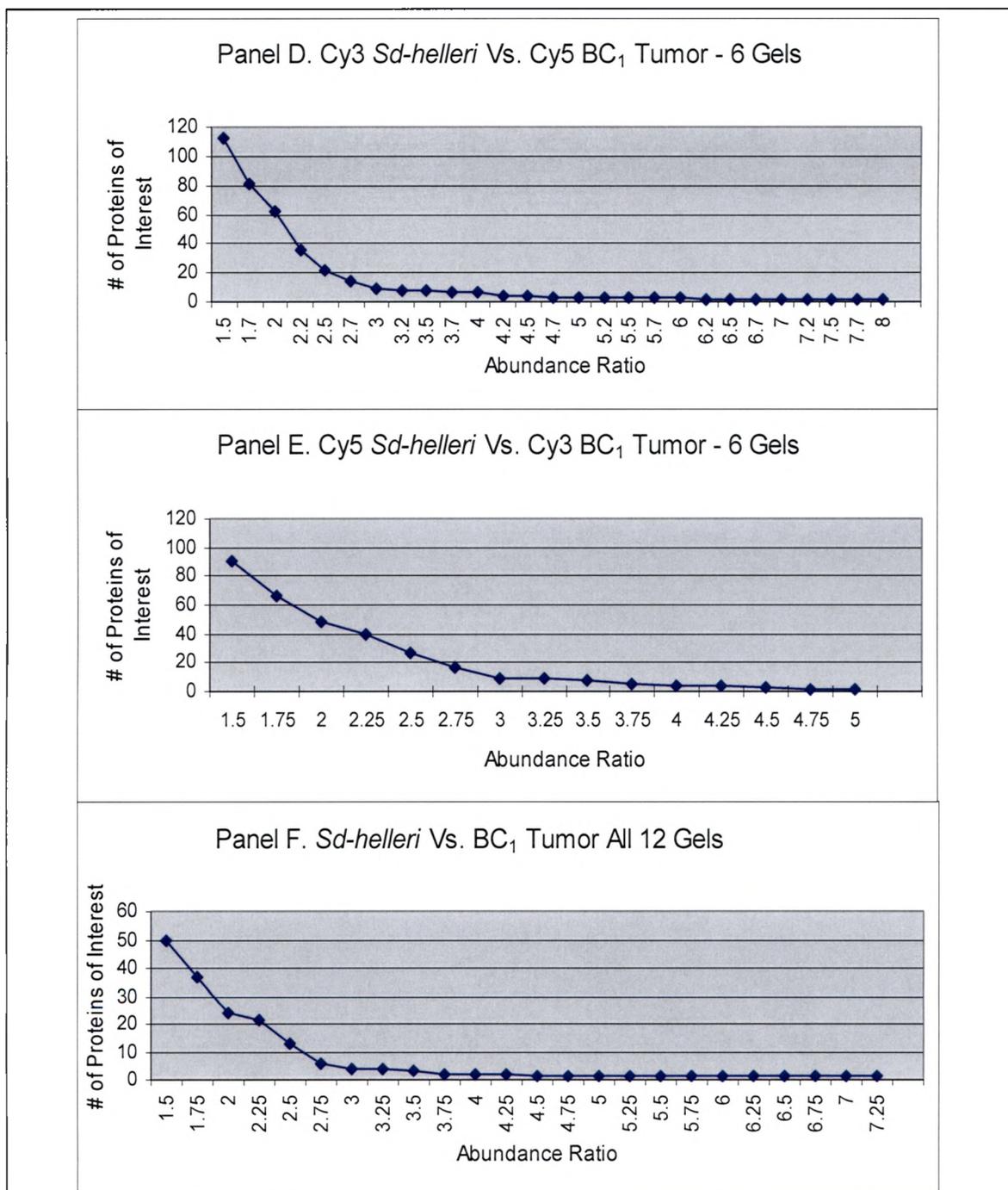


Figure 3-2. Number of proteins of interest versus abundance ratio for three analyses comparing dorsal fin proteins *Sd-helleri* versus BC₁ tumor. Panel D represents an analysis of 6 replicate gels of Cy3-labeled *Sd-helleri* and Cy5-labeled BC₁ tumor. One hundred and thirteen POIs had ± 1.5 fold abundance ratios with the largest ratio detected of 8.24. Panel E represents an analysis of 6 replicate gels of Cy5-labeled *Sd-helleri* and Cy3-labeled BC₁ tumor. Ninety-one POIs had ± 1.5 fold abundance ratios with the largest ratio detected of 5.05. Panel F represents an analysis of 12 replicate gels including 6 gels each from both dye labeling schemes (panels D and E). Fifty POIs had ± 1.5 fold abundance ratios with the largest ratio detected at 7.30.

2) Summary of Protein Spot Data Statistics

Table 3-1 contains protein spot data statistics of the analyses comparing dorsal fin proteins of *X. maculatus* vs. *Sd-helleri* and the analyses comparing *Sd-helleri* vs. BC₁ tumor.

Table 3-1. Protein spot data for DIGE analyses comparing parental <i>X. maculatus</i> , <i>Sd-helleri</i> F ₁ interspecies hybrids, and melanoma tumor from BC ₁ hybrids. All protein samples were from dorsal fin protein fractionated on pH 3-7 DIGE gels.					
1 Dye labeling Scheme ^a		2 # of Gels Analyzed ^b	3 Avg. # Spots Detected ^c	4 Spots Matched between all Gels ^d	5 Proteins of Interest ^e
Cy3	Cy5				
A. F ₁	vs. X mac.	6	879	113 (13 %)	35 (31 %)
B. X. mac	vs. F ₁	6	794	200 (25 %)	68 (34 %)
C. F ₁	vs. X mac.	12	615	65 (11 %)	23 (35 %)
D. F ₁	vs. BC ₁	6	1250	233 (19%)	113 (49 %)
E. BC ₁	vs. F ₁	6	1258	202 (16%)	91 (45 %)
F. F ₁	vs. BC ₁	12	1254	107 (9 %)	50 (47 %)
^a CyDye (3 or 5) used to fluorescently label each sample					
^b Number of analytical DIGE gels used for each analysis					
^c Average calculated from number of protein spots detected in each DIGE gel					
^d Protein spots matched between all DIGE gels in each analysis. In parenthesis, the number represents the percentage of spots matched out of average spots detected					
^e Protein spots that were detected as significant with a t-test α score ≤ 0.05 , +/- 1.5 fold difference abundance ratio, and with 80% power. Number in parenthesis represents the percentage of the number proteins of interest out of the number of matched protein spots.					

The average number of spots detected (table 3-1, column 3) in the gels between all three analyses (sets A, B, and C), *X. maculatus* versus *Sd-helleri*, varied by 264 protein spots. However, the average number of protein spots detected (table 3-1, column 3) in the three analyses comparing *Sd-helleri* versus BC₁ tumor tissue (sets D, E, and F) varied by 8 protein spots. The average number of protein spots detected in analyses A, B and C were 1/3 less than the

average number of protein spots detected in analyses D, E, and F. Analyses A and B are composed of six gels each comparing *X. maculatus* vs. *Sd-helleri* with opposite dye labeling schemes. For analyses A and B differed by 87 protein spots that matched in all gels of both analyses. Analyses D and E are composed of six gels each comparing *Sd-helleri* versus BC₁ tumor with opposite dye labeling schemes. Analyses D and E differed by 31 protein spots that matched in all gels of both analyses. Interestingly, analyses B, D, and E all had around 200 matched protein spots in all gels of each analysis (200, 233 and 202, respectively) where as analysis A had only 113 POIs. In analyses C and F, the number of spots matched in the 12 gels in each analysis compared to the average number of spots detected is 11% and 9%. In the three analyses A, B, and C, the average percentage of significant protein spots (table 3-1, column 5) out of spots matched (table 3-1, column 4) were 33% in contrast to the average percentage of 47% in the three analyses D, E, and F.

3) Validation of Significant Spots Using Power Equation

In Table 3-2, the validation results are reported for the three analyses (A, B, and C) comparing dorsal fin proteins of *X. maculatus* and *Sd-helleri*. An average coefficient of variation was determined for each dye, Cy3, Cy5, and Cy2. Cy3 and Cy5 each represent a sample (*X. maculatus* or *Sd-helleri*). Cy2 represents a combined sample of equal amounts of *X. maculatus* and *Sd-helleri* that was used in the DeCyder analysis as an internal standard. The average coefficient of variation for each dye was used in equation 1. Table 3-2 is a

summary of our findings when the average CV for each dye was used in equation 1 for each analysis. One factor manipulated in equation 1 was the gel replicate number. The gel replicate number represents how many gels needed to be run in order to have significant POIs when the abundance ratio of +/-1.5 fold is held constant. The other factor that can be manipulated in equation 1 was the abundance ratio. When the gel replicate number is held steady (6 or 12 depending on the analysis), we determined what abundance ratio would be considered statistically significant with 80% power.

Table 3-2, analysis C, represents 12 replicate gels with six gels from each opposite dye labeling schemes comparing *X. maculatus* versus *Sd-helleri*. When the coefficients of variation for analysis C is used, the average number of gels needed for the abundance ratios of +/-1.5 to be considered statistically significant with 80% power was 12.3 gels. In analysis C, the statistically significant average abundance ratio of +/-1.5 was determined as based on the set number of 12 replicate gels. The number of gels needed to be run with a constant +/-1.5 fold abundance ratio in analysis B was 13.2 gels. With the increased variation observed in analysis A, holding the abundance ratio steady at a 1.5 fold difference required that 17.5 gels be run. The average abundance ratio needed to maintain 80% power in analysis B using the six replicate gels is 1.89 which is lower than the abundance ratio needed in analysis A of 2.09.

Table 3-2. Validation results for <i>X. maculatus</i> versus <i>Sd-helleri</i> F ₁ hybrid with $\alpha = 0.05$ and 80% power.			
A. 6 Gels - Cy5 <i>X. maculatus</i> / Cy3 <i>Sd-helleri</i>			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.406	16.7	2.05
<i>X. Mac</i>	0.435	19.0	2.16
<i>Sd-helleri</i>	0.408	16.9	2.06
B. 6 Gels - Cy3 <i>X. maculatus</i> / Cy5 <i>Sd-helleri</i>			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.356	13.1	1.88
<i>X. Mac</i>	0.352	12.9	1.88
<i>Sd-helleri</i>	0.363	13.6	1.90
C. All 12 Gels <i>X. maculatus</i> Versus <i>Sd-helleri</i>			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.336	11.8	1.49
<i>X. mac</i>	0.339	12.0	1.50
<i>Sd-helleri</i>	0.351	13.0	1.52
^a Protein sample differentially labeled with Cy2, Cy3, or Cy5			
^b Determined by averaging all coefficients of variation (CV) from all spots in an analysis that were matched in all gels For each spot, CV = standard deviation / average spot volume			
^c In equation 1, $N = 1 + 2C(CV/D)$, N is the number of gels which is solved for when the abundance ratio of 1.5 fold is held constant.			
^d In equation 1, $N = 1 + 2C(CV/D)$, N is the number of gels which is held constant while solving for the abundance ratio represented by D. $D = \ln(\text{abundance ratio})$			
^e Cy2 is the internal standard used in DIGE. Cy2 labels equal amounts of <i>X. maculatus</i> and <i>Sd-helleri</i> .			

Table 3-3. Validation results for <i>Sd-helleri</i> versus BC ₁ hybrid tumor tissue with $\alpha = 0.05$ and 80% power.			
D. 6 Gels - Cy3 <i>Sd-helleri</i> / Cy5 BC ₁ Tumor			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.405	16.7	2.05
<i>Sd-helleri</i>	0.415	17.4	2.09
BC ₁ Tumor	0.401	16.4	2.04
E. 6 Gels - Cy5 <i>Sd-helleri</i> / Cy3 BC ₁ Tumor			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.457	20.9	2.25
<i>Sd-helleri</i>	0.486	23.6	2.37
BC ₁ Tumor	0.465	21.6	2.28
F. All 12 Gels <i>Sd-helleri</i> Versus BC ₁ Tumor			
Protein Sample ^a	Coefficient of Variation ^b	Gels Needed to Run with 1.5 Ratio ^c	Abundance Ratio with Current Gel Count ^d
Cy2 ^e	0.465	21.6	1.74
<i>Sd-helleri</i>	0.504	25.3	1.83
BC ₁ Tumor	0.478	22.8	1.77
^a Protein sample differentially labeled with Cy2, Cy3, or Cy5			
^b Determined by averaging all coefficients of variation (CV) from all spots in an analysis that were matched in all gels. For each spot, CV = standard deviation / average spot volume			
^c In equation 1, $N = 1 + 2C(CV/D)$, N is the number of gels which is solved for when the abundance ratio of 1.5 fold is held constant			
^d In equation 1, $N = 1 + 2C(CV/D)$, N is the number of gels which is held constant while solving for the abundance ratio represented by D. $D = \ln(\text{abundance ratio})$			
^e Cy2 is the internal standard used in DIGE. Cy2 labels equal amounts of <i>Sd-helleri</i> and BC ₁ tumor.			

Table 3-3 presents an analysis (F) of 12 replicate gels with six gels in each dye labeling direction that compares dorsal fin proteins of *Sd-helleri* versus BC₁ tumor. When the CV for analysis F was used in equation 1, the average number of gels needed for the abundance ratio of +/-1.5 to be statistically

significant with 80% power was 23.2 gels. In analysis F, the statistically significant average abundance ratio of 1.78 was determined based on the set number of 12 replicate gels. The number of gels that would be needed to run with a constant +/-1.5 fold abundance ratio in analysis D was 16.8 gels. With the increased variation seen in analysis E, holding the abundance ratio steady at a +/-1.5 fold difference required that 20.0 gels be run. The average abundance ratio needed to maintain 80% power in analysis D using the six replicate gels is 2.06 which is lower than the average abundance ratio needed in analysis E of 2.30.

D. DISCUSSION

Comparison of dorsal fin proteins of the parental *X. maculatus* and the *Sd-helleri* F₁ interspecies hybrid was conducted in three separate analyses, A, B, and C. When the abundance ratio requirement was increased to a 2.0 fold difference, each of the three analyses exhibited decreases to about half the initial number of POIs compared to the abundance ratios of +/-1.5 fold that was initially used. In table 3-1, analysis A, the number of POIs are about half the amount of the POIs seen in the dye flip, analysis B. This difference in POI number appears to be due to analysis of experiment A indicating the total protein spots matched in all gels were half the amount of the protein spots able to be matched in dye flip gels in analysis B. The analyses comparing *X. maculatus* versus *Sd-helleri* (A, B, and C) had 1/3 less protein spots detected by the software than the number of protein spots detected in the analyses comparing *Sd-helleri* versus BC₁ tumor (D,

E, and F). The lower amount of detectable protein spots could be due to technical variation that may occur in the analytical process; including but not limited to labeling inconsistencies, gel running conditions (IEF gel conditions and SDS-PAGE), and variations in matching during the image analysis (25). Interestingly, despite varying amounts of proteins that were matched in all gels among the three analysis (A = 113, B = 200 and C = 65), the percentage of the matched spots found to be statistically significant were all in the lower 30% range. This may suggest that 30% of all proteins in the F₁ hybrid dorsal fin tissue see significant differences in abundance when compared to proteins of the parent *X. maculatus*.

In table 3-1, the differences in the number of protein spots detected between the three analyses (D, E, and F) comparing *Sd-helleri* to BC₁ tumor were within 8 protein spots of each other. The protein spots matched in all gels in analysis D and the opposite dye labeling scheme in analysis E were different by 31 protein spots. The 31 protein spot difference could be due to preferential labeling by one dye over another depending on the properties of the individual proteins (24).

Despite the inconsistencies in spot matching numbers between labeling assignments in the *X. maculatus* versus *Sd-helleri* comparison, the observed variation represented by the CV appears to be compensated for by running 12 replicate gels. In this case the 12 gels resulted in statistically significant abundance ratios at +/-1.5 fold or larger between differentially labeled samples. Analyses A and B had similar CVs to that of analysis C. However, the abundance

ratio had to be increased to an average of 1.98 to compensate for the fewer number of repeats. Karp *et al.* (29), observed that variance tended to be largest in the low abundance spots. Filtering these proteins out of the analysis may, however, not prove to be the best answer. The proteins that are in low abundance are more likely to provide drug targets or diagnostic markers and their abundance changes may provide more insight for discovery (34). Our analyses indicate the number of needed replicate gels was met in the 12 gel comparison of the dorsal fins of *X. maculatus* versus *Sd-helleri* allowing us to keep the initially-used abundance ratio levels at +/-1.5 fold.

The analyses comparing fin proteins of *Sd-helleri* and BC₁ tumor (D, E, and F) had higher CVs than the previous comparison of *X. maculatus* and *Sd-helleri*. In order to keep the abundance ratio level statistically significant at a +/- 1.5 fold difference between differentially labeled samples, the CV in analysis F required an average of 24 replicate gels which doubles from the 12 replicate gels ran. Adjusting the statistically significant abundance level in analysis F to approximately a 1.75 fold change in protein abundance allowed the 12 replicate gel number to remain stable while resulting in a loss of 13 out of the initial 50 POIs. One hypothesis for the cause of increased variability in these three analyses (D, E, and F) is the multitude of cell types in melanoma tumor including normal, necrotic, and dead cells. The two six gel analyses (D and E) had small decreases in their CV in comparison to analysis F. Even though the CVs decreased, the abundance ratios needed for statistical significance increased to a 2.0 fold difference with only 6 replicate gels. Thus, illustrates the fine balance

between the number of replicates needed for a target abundance ratio to compensate for CV and still detect statistically significant POIs.

Having a better understanding of the protein spot volume variability from gel to gel and accommodating for Type I and Type II errors through the validation process (i.e. equation 1) increases our confidence the POIs chosen are significant. The DeCyder analysis validates the protein spots and considers Type I error with t-tests conducted on the protein spots. Equation 1 has been adapted to consider protein spot volume variation from gel to gel and its effects on Type II error confidence. This validation procedure insures added certainty that the abundance level and replicate number used in our experimental design provides truly significant POIs for further analyses.

CHAPTER IV

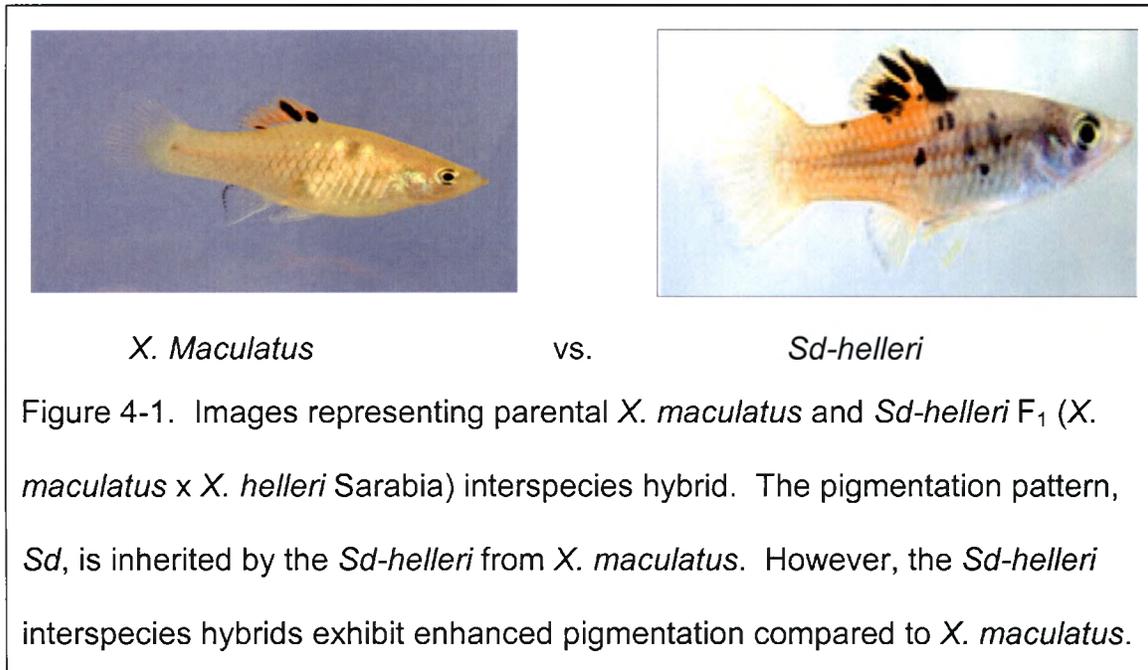
DIGE COMPARISON OF DORSAL FIN PROTEINS FROM *Xiphophorus maculatus* AND *Sd-helleri* F₁ INTERSPECIES HYBRIDS

A. INTRODUCTION

In order to identify proteins exhibiting altered abundance due to the development of melanoma, we must first identify proteins that exhibit altered abundance due solely to combining genomes by interspecies hybridization. In this DIGE comparison, we examined protein differences between the dorsal fins of parental, *X. maculatus* Jp 163 A, and *Sd-helleri* F₁ hybrid progeny resulting from crossing this strain with *X. helleri* (Sarabia) (Figure 4-1).

B. METHODS

For this comparison of *X. maculatus* to *Sd-helleri*, a 50 µg dorsal fin protein sample from *X. maculatus* and a 50 µg dorsal fin protein sample from *Sd-helleri* was labeled with 200 pmol of cyanine-3-fluorescent dye (Cy3) and Cy5, respectively. This dye labeling scheme was repeated in one repeat labeling. The Cy3 and Cy5 labeling assignments were reversed in two subsequent labelings as "dye flips". A control sample combining 25 µg of dorsal fin protein sample from



X. maculatus and 25 µg of dorsal fin protein sample from *Sd-helleri* was labeled with Cy2 for use as an internal standard for each gel run. A gel run consists of four analytical DIGE gels including two replicate gels of Cy3-labeled *X. maculatus* and Cy5-labeled *Sd-helleri* and two replicate gels of Cy5-labeled *X. maculatus* and Cy3-labeled *Sd-helleri*. The dorsal fin proteome comparison of *X. maculatus* versus *Sd-helleri* was conducted in 12 replicate gels or three replicate gel runs.

A Visual analysis method was used that split up the 12 replicate gels into 6 sets of two gels each. These two gels were replicates of the same dye labeling scheme of the same gel run. An analysis was conducted on each of these six sets of gels. Significant protein spots in each of the six analyses had to meet the following criteria of the DeCyder software: (1) protein spots had to exhibit abundance ratios starting at 1.5 fold or greater between differentially labeled samples; (2) protein spots had to have a t-test score of $\alpha \leq 0.05$. If a protein

spot in these analyses met these two criteria they were termed a protein of interest (POI). Then a representative gel image was selected from each analysis with their POIs located. The gel images were compared visually for common POIs. The POIs that were common across all six analyses were the POIs chosen for identification by MALDI-TOF MS.

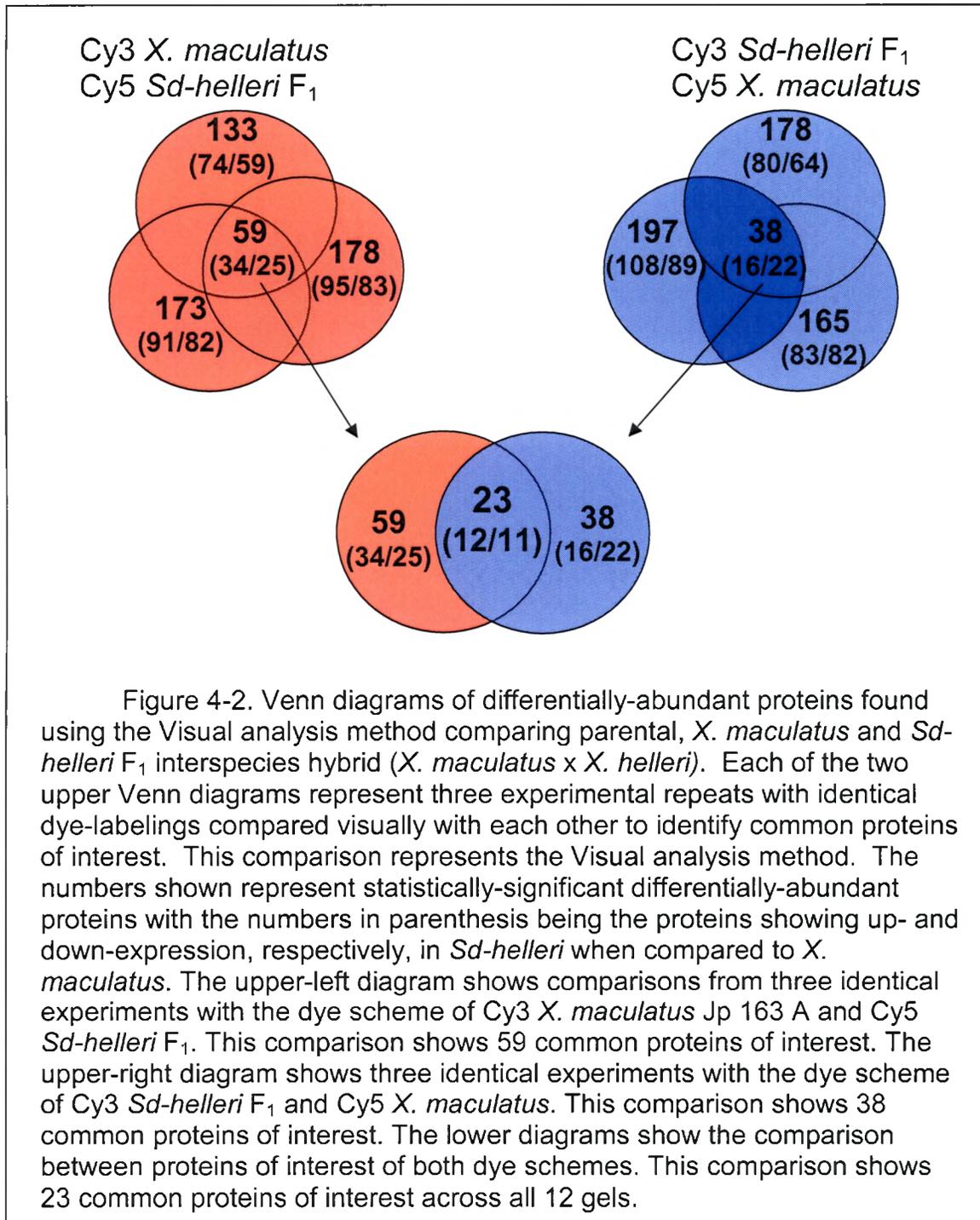
A Validation analysis method was conducted with one analysis using the same 12 replicate gels as the Visual method. The POIs were chosen by the DeCyder software according to the two criteria stated above. In addition, the Validation analysis method analyzed the DIGE gels according to the requirements specified in Chapter 3. The Validation method POIs were chosen with the additional consideration of variation and the requirement of statistical significance with 80% power.

C. RESULTS

1) Venn Diagrams for Determination of Proteins of Interest in 12 Replicate Gels

Figure 4-2 contains Venn diagrams that show how the final proteins of interest were identified by the Visual analysis method from six separate analyses containing two gel replicates each.

Each of the three top red circles represent analyses consisting of two replicate gels with the dye scheme of Cy3-labeled *X. maculatus* and Cy5-labeled *Sd-helleri*. The POIs observed in each analysis ranged from 133 to 178 POIs. The analyses were visually compared to find the POIs that were common between the three analyses. Fifty-nine proteins were found to be common



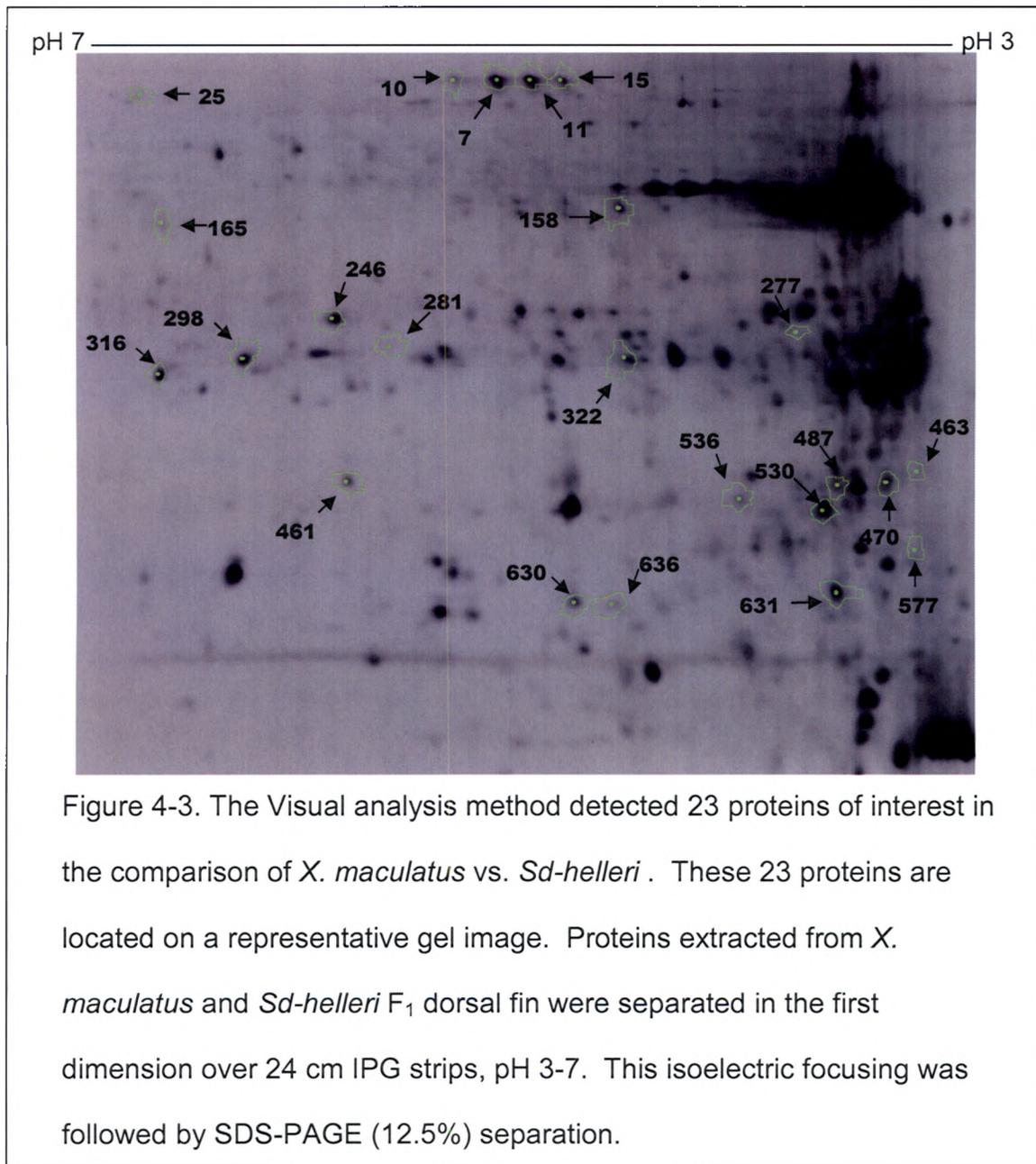
between the three analyses with 34 proteins showing higher abundance in *Sd-helleri* and 25 showing lower abundance in *Sd-helleri*.

Each of the three top blue circles represent analyses consisting of two replicate gels with the dye labeling scheme of Cy5-labeled *X. maculatus* and Cy3-labeled *Sd-helleri*. POIs observed in the analyses ranged from 165 to 197 POIs. These three analyses were also visually compared to find common POIs. Thirty-eight proteins were common between the three analyses with 16 proteins exhibiting higher abundance in *Sd-helleri* and 22 proteins exhibiting lower abundance in *Sd-helleri*. The 59 POIs of the first dye scheme and the 38 POIs of the opposite dye labeling scheme were visually analyzed for common POIs. Twenty-three POIs were common in all 12 gels represented across the two dye labeling schemes with 12 POIs exhibiting higher abundance in *Sd-helleri* and 11 POIs exhibiting lower abundance in *Sd-helleri*.

2) Gel Images Representing Proteins of Interest Identified by the Visual and Validation Analysis Methods

Figure 4-3 is a representative gel image showing the 23 POIs detected using the Visual analysis method. The POIs are spread across the entire gel with some clustering at the top of the gel and in the lower right hand corner.

Figure 4-4 is a representative gel image of the 23 POIs detected using the Validation analysis method. There are less POIs clustered at the top of this gel in comparison to the POIs in the Visual method gel but the clustering of POIs at the lower right hand corner is similar.



When comparing the gel images of the Visual and the Validation method, 8 POIs were found to be common. Table 4-1 gives the identification numbers of the 8 common POIs from figure 4-3 detected by the Visual analysis method and

the corresponding identification numbers from figure 4-4 detected by the Validation analysis method. Two common protein pairs have identical abundance ratios from both methods (298/556 and 631/944).

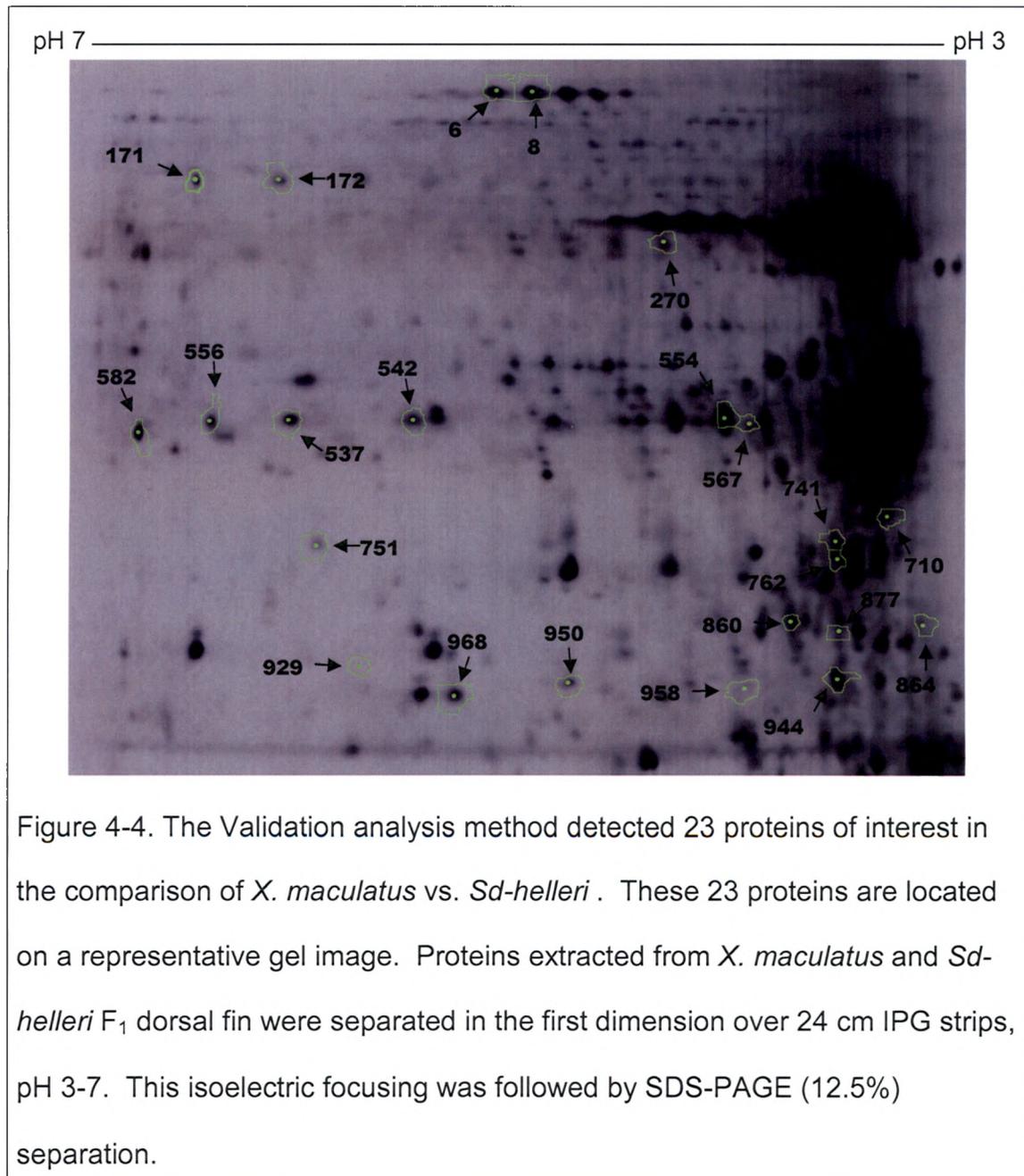


Figure 4-4. The Validation analysis method detected 23 proteins of interest in the comparison of *X. maculatus* vs. *Sd-helleri*. These 23 proteins are located on a representative gel image. Proteins extracted from *X. maculatus* and *Sd-helleri* F₁ dorsal fin were separated in the first dimension over 24 cm IPG strips, pH 3-7. This isoelectric focusing was followed by SDS-PAGE (12.5%) separation.

Four additional protein pairs have abundance ratios within 0.05 of one another. There are two protein pairs whose abundance ratios differ by 0.39 between analysis methods.

Table 4-2 lists the Validation method POIs according to their identification numbers assigned in figure 4-4 and their corresponding abundance ratios. Fourteen POIs showed statistically significant lower abundance ratios in *Sd-helleri* whereas 9 proteins showed statistically significant higher abundance ratios in *Sd-helleri*.

Table 4-1. Corresponding identification numbers for common proteins of interest between the Visual and Validation methods presented in figure 4-3 and figure 4-4.				
Figure 4-3 ID # ^a	Abundance Ratio ^b		Abundance Ratio ^b	Figure 4-4 ID # ^c
7	-2.75		-2.80	6
11	-2.16		-2.20	8
298	1.95		1.95	556
316	-1.88		-1.84	582
461	1.60		1.56	751
487	-2.27		-1.88	762
630	-2.23		-1.84	950
631	1.68		1.68	944

^a Protein spot number corresponding to POI number assigned in figure 4-3 resulting from the Visual analysis method

^b Abundance ratios of each protein of interest as seen in *Sd-helleri* in comparison to *X. maculatus*

^c Protein spot number corresponding to POI number assigned in figure 4-4 resulting from the Validation analysis method

Table 4-2. Proteins of interest identified in figure 4-4 using the Validation analysis method.

Figure 4-4 ID # ^a	Abundance Ratio ^b
929	3.81
864	2.23
958	2.01
556	1.95
944	1.68
567	1.64
537	1.61
751	1.56
542	1.53
171	-1.50
554	-1.59
172	-1.60
968	-1.64
270	-1.78
582	-1.84
950	-1.84
710	-1.87
762	-1.88
877	-1.88
860	-2.02
8	-2.22
741	-2.39
6	-2.80

^a Protein spot number corresponding to POI number assigned in Figure 4-4

^b Abundance ratios of each protein of interest as seen in *Sd-helleri* in comparison to *X. maculatus*

3) Protein Identifications of Proteins of Interest from the Visual Analysis

The Validation method was not completed until after the protein identification process was conducted

on the POIs detected using the Visual method. Therefore, the identification of the POIs detected using the Validation method is in progress.

Table 4-3 presents the 23 POIs determined by the Visual analysis method by their identification number, their corresponding abundance ratios in *Sd-helleri* compared to *X. maculatus*, the weights of parental ions that were fragmented using MALDI-TOF MS, the amino acid sequences obtained by postsource decay fragmentation, and possible protein identifications indicated from searching the amino acid sequences against public protein databases.

Of the 23 proteins, 5 POIs were unable to provide peptides that resulted in amino acid sequences.

Amino acid sequences were obtained for the remaining 18 POIs with varying results in identification confidence.

An expect value is a statistical expectation value which controls the level of similarity needed for a match to be reported (35). In general, sequences with expect values of less than 0.01 are homologous and expect values between 1 and 10 indicate relation to the protein target resulting in possible identification (35). Fourteen proteins were identified with an expect value of less than 1. Four proteins were identified as being possibly related to known protein families with an expect value of less than 10. Proteins identified with the best expect value scores included transferrin ($E = 2 \times 10^{-04}$), actin ($E = 8 \times 10^{-06}$), peroxiredoxin ($E = 2 \times 10^{-05}$), keratin ($E = 5 \times 10^{-04}$), and glutathione-S-transferase ($E = 0.013$).

D. DISCUSSION

The DIGE comparison of parental *X. maculatus* and the *Sd-helleri* F₁ interspecies hybrid (*X. maculatus* x *X. helleri*) provides a baseline for protein abundance alterations due solely to genetic hybridization. The DIGE Visual analysis method detected 23 POIs as did the Validation analysis method. However, when Visual and Validation POI's locations on their respective gel images were compared, we found only 8 POIs in common. Identification of all POIs from the Validation analysis is in progress. Still, the proteins identified from the Visual analysis initiate our understanding of the effects interspecies hybridization may have on basal molecular genetic and biochemical regulation.

In order to classify proteins as melanoma biomarkers, we first have to identify proteins that are affected due solely to the hybridization of two distinct species. Initially, a Visual analysis method was utilized that separated the 12 replicate gels comparing *X. maculatus* to *Sd-helleri* and separated them into 6 DIGE analyses. The six analyses detected as many as 197 POIs and as few as 133 POIs . Each of the six analyses generated a representative gel image with the POIs located on it. These six gel images were visually compared for common POIs. As a result, 23 POIs were found to be common between all six analyses. The reduction from the highest count of 197 POIs in one analysis to 23 POIs found to be common between all six analyses is a drastic decline that clearly shows the need for reproducibility across repeats and the difficulty in obtaining exact experimental replication. Reproducibility is traditionally the principal downfall of DIGE and other 2-D gel techniques (19).

The Validation analysis identified an identical number of POIs as the Visual analysis method. However, the Validation analysis method provided additional requirements for proteins to be termed POIs. Two advantages of the Validation analysis method include better t-test values and the ability to apply consistent validation requirements. The better t-test values are due to the increased number of replicates (30). The validation requirements would need to be implemented in either case (12 gels in one analysis or six analyses of 2 gels each) but having all data contained in one analysis set will accommodate for variation in all gels across both dye schemes. Looking at variation in all the gels in a single analysis will give a better representation of the average variation seen

within an experiment. Thus, the proteins selected as POIs will be based upon statistically significant confidence intervals. These proteins are ones that we wish to focus on for identification and insight to the interspecies hybridization process.

One of the proteins chosen for amino acid sequencing and identification from the Visual method was transferrin. Transferrin mediates iron levels in the cell (36). Kang *et al.* (2005) found that a reduction of transferrin levels can induce apoptosis by a sodium ascorbate-mediated pathway. Transferrin is interesting because of its multiple effects on the cell including metabolism and cellular longevity.

Another important protein identified was glutathione-S-transferase (GST). The GST gene has been implicated in familial melanoma cases (38). GST is an antioxidant protein that exhibits higher levels of activity in skin cancer thought to be caused by the increased oxygen free radicals produced in skin cancer (39).

Peroxiredoxin is a redox-active antioxidant (40). Peroxiredoxin was found to be almost two times higher in *Sd-helleri* than in *X. maculatus*. Peroxiredoxin is known for its antioxidant properties but is also important for redox dependent cell signaling (40). A future question to address is why the abundances of antioxidants like peroxiredoxin and GST are affected in opposite ways in the dorsal fin cells of *Sd-helleri* compared to that of the parental, *X. maculatus*? Could the preference for peroxiredoxin be due to the type of oxidation affecting the cell? Is peroxiredoxin more useful to the cell in addressing these oxidation issues than GST?

POIs detected by the Visual method that were identified have very different functions but provide us with possible cellular functions that may be affected due to hybridization alone. Spontaneous melanoma develops in the progeny resulting from the backcross hybridization between *Sd-helleri* and *X. helleri*. This study of the protein differences between *X. maculatus* and *Sd-helleri* lays the foundation for identifying which cellular changes observed in melanoma, are not due to the hybridization but unique protein changes due to cellular changes occurring between melanized pigment cells and melanoma development.

Table 4-3. Protein identifications of *X. maculatus* Vs. *Sd-helleri* F₁ DIGE proteins of interest.

Protein of Interest ^a	Sd-helleri F ₁ Abundance Ratio ^b	SPITC-peptide precursor ion (m/z) ^c	Peptide Sequence ^d	Protein Identification ^e
7	-2.75	1897.438	EQYYGYAGAFR	Transferrin-a (<i>D. rerio</i> , E = 0.002)
		1898.520	EEGYGYAGAFR	Transferrin-a (<i>D. rerio</i> , E = 2 x 10 ⁻⁰⁴)
165	-2.27	1940.946	GTQENGE[I/L]VD[I/L]AAFSR	Phospholipase AdRab-B precursor (<i>O. cuniculus</i> , E = 0.084)
630	-2.23	1164.383	E[I/L]GVGFATR	Fatty acid binding protein (<i>M. musculus</i> , E = 0.38)
246	-2.20	2300.006	P[I/L]GQEVQ[I/L]S[I/L][I/L]PR	Peroxisome assembly factor (<i>R. norvegicus</i> , E = 0.032)
		1339.423	[I/L]QYCD[I/L]VDR	Protocadherin cluster 2 gamma 4 (<i>D. rerio</i> , E = 3.8)
		1646.890	AQDGFVND[I/L]A[I/L][I/L]R	Similar to molybdenum cofactor - step one protein isoform 3 (<i>D. rerio</i> , E = .31)
11	-2.16	1898.230	EEGYGYAGAFR	Transferrin-a (<i>D. rerio</i> , E = 2 x 10 ⁻⁰⁴)
10	-2.11	1178.629	VPAHAV[I/L]TR	Serotransferrin precursor (<i>X. laevis</i> , E = 0.002)
		1898.430	YYGYAGAFR	Transferrin-a (<i>D. rerio</i> , E = 0.20)
158	-2.10	1606.250	EQVTY[I/L]Q[I/L][I/L]ER	Plectin (<i>R. norvegicus</i> , E = 0.013)
15	-1.88	1178.447	VPAHAV[I/L]TR	Precursor to transferrin-a (<i>D. rerio</i> , E = 9.2)
		1922.940	QYYGYAQFR	Transferrin-a (<i>D. rerio</i> , E = 0.38)
316	-1.88	1334.457	V[I/L][I/L]HYFDGR	Glutathione S-transferase (<i>O. cuniculus</i> , E = 0.013)
		1272.467	A[I/L][I/L]HY[I/L]DGR	Glutathione S-transferase (<i>O. cuniculus</i> , E = 0.013)
		1231.445	MTQ[I/L]PA[I/L]SR	Acetyltransferase (<i>E. coli</i> , E = 0.35)
461	1.60	1968.572	PDG[I/L]ED[I/L][I/L]GSR	Phosphoinositide 3 kinase catalytic subunit (<i>X. laevis</i> , E = 6.5)
		1552.275	[I/L]EYTADYSPY	Putative mitochondrial protein (<i>C. elegans</i> , E = 6.5)
636	1.65	1344.625	AF[I/L]FAEY[I/L]R	Novel zinc finger protein (<i>D. rerio</i> , E = 5.1)
		1708.757	SDPGPTYN[I/L]T[I/L]GR	Transmembrane channel-like protein (<i>H. sapiens</i> , E = 0.24)

631	1.68	1730.620	[I/L]WHHTFYNE[I/L]R	Actin (<i>D. rerio</i> , E = 1 x 10 ⁻⁰⁴)
		1160.615	AVFPS[I/L]VGR	Actin (<i>G. gallus</i> , E = 0.043)
298	1.95	2719.200	DEAGTP[I/L]PFP[I/L][I/L]ADDQR	Peroxiredoxin (<i>D. rerio</i> , E = 0.035)
		2334.970	PY[I/L]NGEVFNPFADTTSGR	Peroxiredoxin (<i>D. rerio</i> , E = 0.96)
		1816.695	E[I/L]SVQ[I/L]GM[I/L]DPDER	Peroxiredoxin (<i>D. rerio</i> , E = 2 x 10 ⁻⁰⁵)
277	2.04	1762.447	FWHHTFYNE[I/L]R	Actin (<i>D. rerio</i> , E = 0.005)
		1730.442	[I/L]DAHHTFYNELR	Actin (<i>D. rerio</i> , E = 0.005)
530	2.10	1958.885	ApTFEAY[I/L]AN[I/L]R	Keratin (<i>D. rerio</i> , E = 0.07)
		1606.250	VDALQDELNFLR	Keratin 4 (<i>D. rerio</i> , E = 5 x 10 ⁻⁰⁴)
281	2.08	2097.862	PQQS(E,G)SA[I/L]VD[I/L]	Tensin (<i>C. elegans</i> , E = 0.4)
322	4.12	2005.831	SYE[I/L]PNGQV[I/L]T[I/L]GNER	Actin (<i>D. rerio</i> , E = 8 x 10 ⁻⁰⁶)
		2178.870	SQAQGVNAEQLTSLR	Maltodextrin phosphorylase (<i>E. coli</i> , E = 1.9)
577	4.45	1589.531	AS[I/L]ADVQNR	Keratin (<i>P. reticulata</i> , E = 17)
Proteins With No Sequence or ID				
470	-2.34			
487	-2.27			
25	-2.22			
463	1.88			
536	1.89			

^a Protein spot number corresponding to POI number assigned in Figure 2.

^b Average abundance ratio across 6 analyses representing 2 replicate gels each.

^c Monoisotopic m/z 2:1 TDC matrix used to collect spectra in positive ion mode.

^d Peptide amino acid sequence determined by postsource decay spectra of SPITC-labeled peptides.

^e Protein identification based on best expect value score from several database searches including NCBI, SwissProt, and FASTA at University of Virginia

CHAPTER V

DIGE COMPARISON OF DORSAL FIN PROTEINS FROM *Sd-helleri* F₁ AND BC₁ MELANOMA

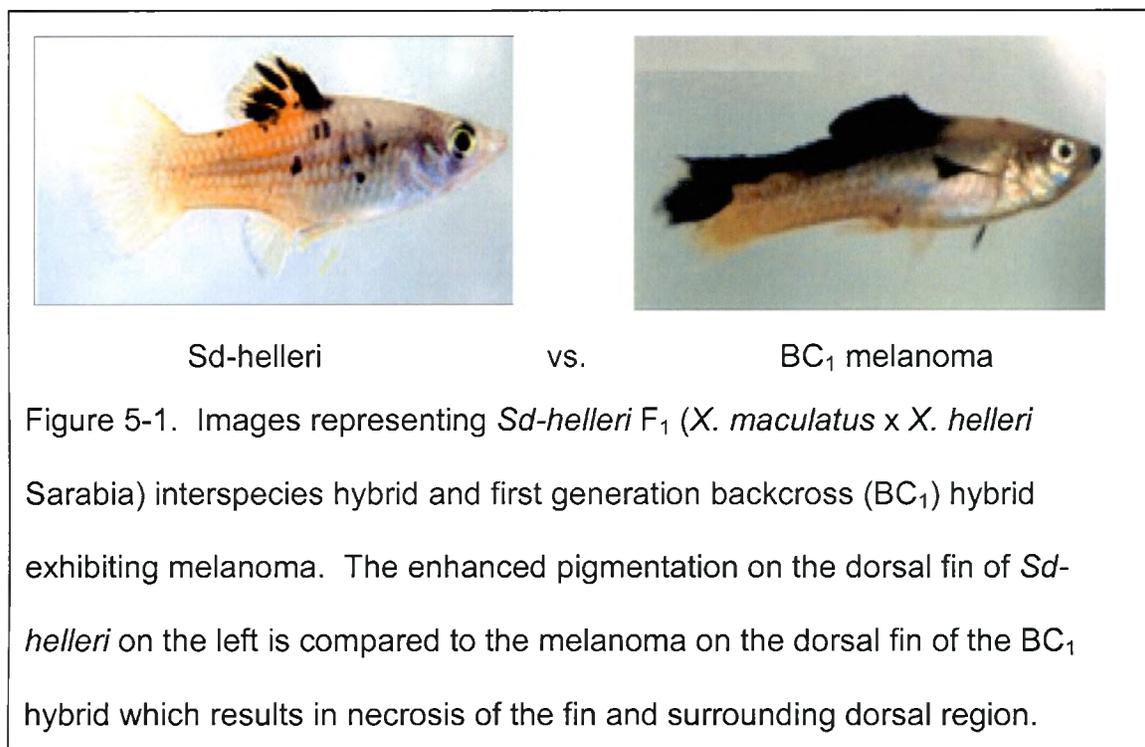
A. INTRODUCTION

Relative protein abundances affected by interspecies hybridization have been examined (see chapter 4). Here we analyze protein abundance differences comparing *Sd-helleri* interspecies F₁ hybrid melanin pigmented dorsal fin tissue and melanoma tissue from first generation backcross (BC₁) hybrid dorsal fins. In this comparison, we will compare and assess protein abundance differences between *Sd-helleri* fin and BC₁ melanoma tissues in order to attempt to identify protein biomarker candidates that are present upon transformation of heavily pigmented cells to melanoma. Figure 5-1 contains images of the fish *Sd-helleri* and the BC₁ hybrid exhibiting melanoma tumor.

B. METHODS

For this comparison of *Sd-helleri* F₁ fin with BC₁ melanoma tissues, 50 µg of dorsal fin protein from *Sd-helleri* and a 50 µg dorsal melanoma protein sample from BC₁ tumor-bearing animals was labeled with 200 pmol of cyanine-3-fluorescent dye (Cy3) and Cy5, respectively. This dye labeling scheme was repeated once. The Cy3 and Cy5 labeling assignments were reversed in two

subsequent labelings as "dye flips". A control sample combining 25 μg of dorsal fin protein sample from *Sd-helleri* F_1 hybrid and 25 μg of dorsal fin protein sample



from melanoma tissue of BC₁ hybrids was labeled with Cy2 for use as an internal standard for each gel run. A gel run consists of four analytical DIGE gels including two replicate gels of Cy3-labeled fin from *Sd-helleri* F_1 hybrids and Cy5-labeled BC₁ melanoma tissue and two replicate gels of Cy5-labeled *Sd-helleri* F_1 hybrid fin tissue and Cy3-labeled BC₁ melanoma. The dorsal fin proteome comparison of *Sd-helleri* F_1 fins versus BC₁ melanoma was conducted in 12 replicate gels or three replicate gel runs.

A Visual analysis method was used that split up the 12 replicate gels into 6 sets of two gels each. These two gels were replicates of the same dye labeling scheme of the same gel run. An analysis was conducted on each of these six sets of gels. As before (see chapter 4) significant protein spots in each of the six

analyses had to meet the following criteria of the DeCyder software: (1) protein spots had to exhibit abundance ratios starting at 1.5 fold or greater between differentially labeled samples; (2) protein spots had to have a t-test score of $\alpha \leq 0.05$. If a protein spot in these analyses met these two criteria they were termed a protein of interest (POI). Then a representative gel image was selected from each analysis with their POIs located. The gel images were compared visually for common POIs. The POIs that were common across all six analyses were the POIs chosen for identification by MALDI-TOF MS.

A Validation analysis method was conducted also applied to the same 12 replicate gels analyzed by the Visual method. The POIs were chosen by the DeCyder software according to the two criteria stated above. In addition, the Validation analysis method analyzed the DIGE gels according to the requirements specified in Chapter 3. The Validation method POIs were chosen with the additional consideration of variation in spot densities between gels and the requirement of statistical significance with 80% power.

C. RESULTS

1) Venn Diagrams Showing Identification of Proteins of Interest

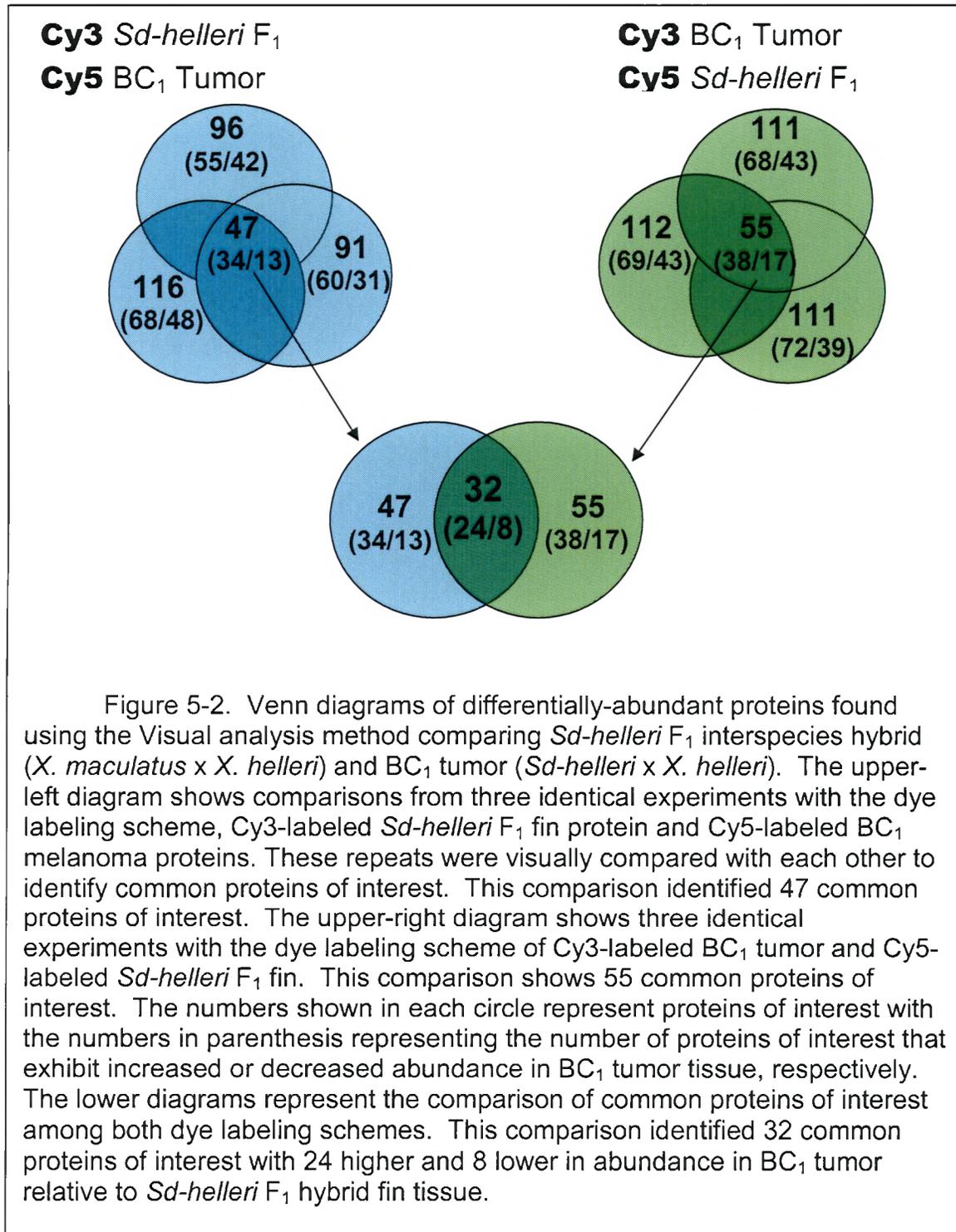
Figure 5-2 presents Venn diagrams that show the final POIs identified by the Visual analysis method from the six separate analyses containing two gel replicates each.

Each of the blue circles in the left-hand corner represent analyses consisting of two gels each with the dye scheme of Cy3-labeled *Sd-helleri* F₁ fin

and Cy5-labeled BC₁ melanoma. The number of POIs observed in these three analyses ranged from 91 to 116 POIs. The analyses were visually compared to find POIs that were common between the three analyses. Forty-seven proteins were found to be common between all three analyses with 34 proteins showing higher abundance in BC₁ melanoma tumor tissue and 13 proteins showing lower abundance in BC₁ melanoma tumor tissue.

Each of the three top green circles in the right-hand corner represent analyses of two gels, each with the dye labeling scheme of Cy3-labeled BC₁ melanoma tissue and Cy5-labeled *Sd-helleri* dorsal fin tissue. The number of POIs detected in the three analyses ranged from 111 to 112 POIs. Fifty-five proteins were found to be common between the three analyses with 34 POIs showing higher abundance in BC₁ tumor and 17 POIs showing lower abundance in BC₁ tumor.

The 47 POIs of the first dye labeling scheme and the 55 POIs of the second dye labeling scheme were visually analyzed for common POIs. Thirty-two proteins were common between all six analyses with 24 POIs exhibiting higher abundance in BC₁ tumor and 8 POIs exhibiting lower abundance in BC₁ tumor.



2) Gel Images Showing Proteins of Interest Identified by Visual and Validation Analysis Methods

Figure 5-3 is a representative gel image showing the 32 POIs detected using the Visual analysis method. The POIs are spread evenly within the pH range (shown at top) with some concentration in the middle of the molecular weight distribution.

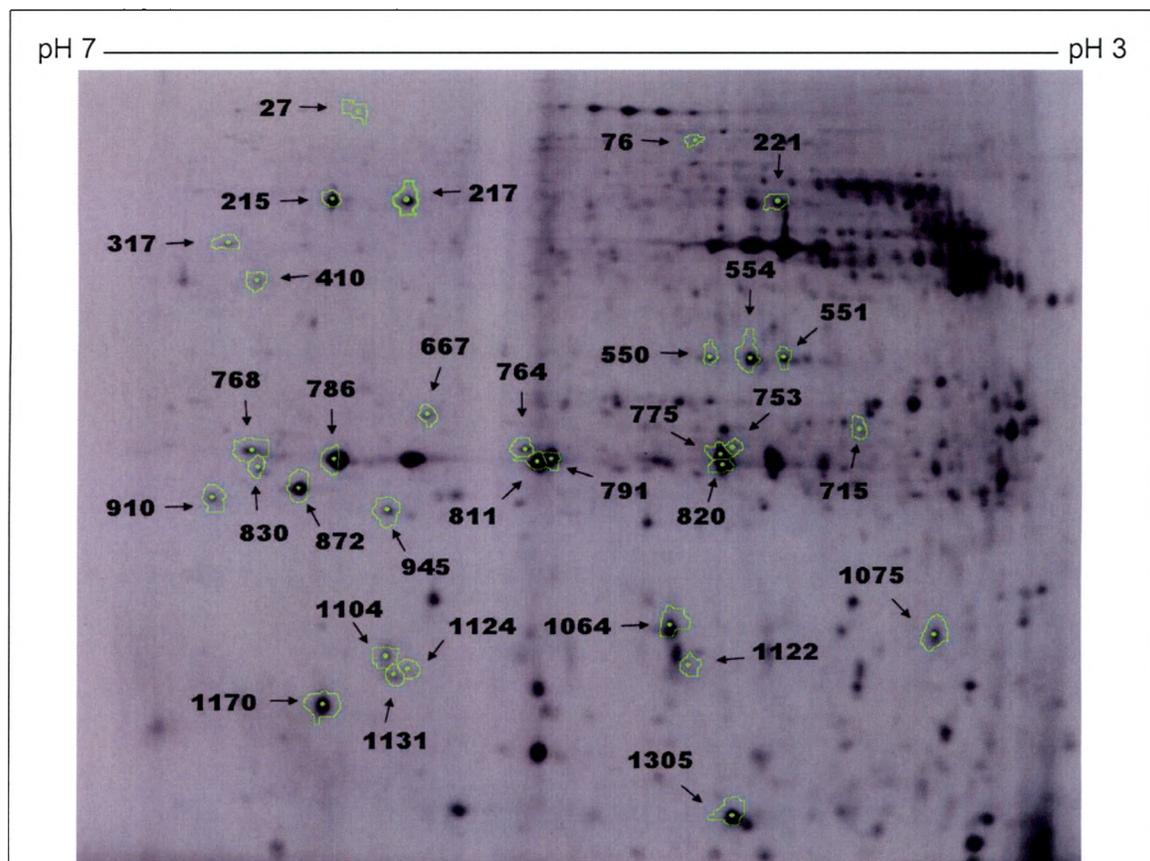


Figure 5-3. The Visual analysis method detected 32 proteins of interest in the comparison of *Sd-helleri* fin vs. BC₁ melanoma tissue. These 32 proteins are shown on a representative gel image (circled in green). Proteins extracted from *Sd-helleri* fin and BC₁ tumor were separated in the first dimension over 24 cm IPG strips, pH 3-7. The isoelectric focusing was followed by SDS-PAGE (12%) separation.

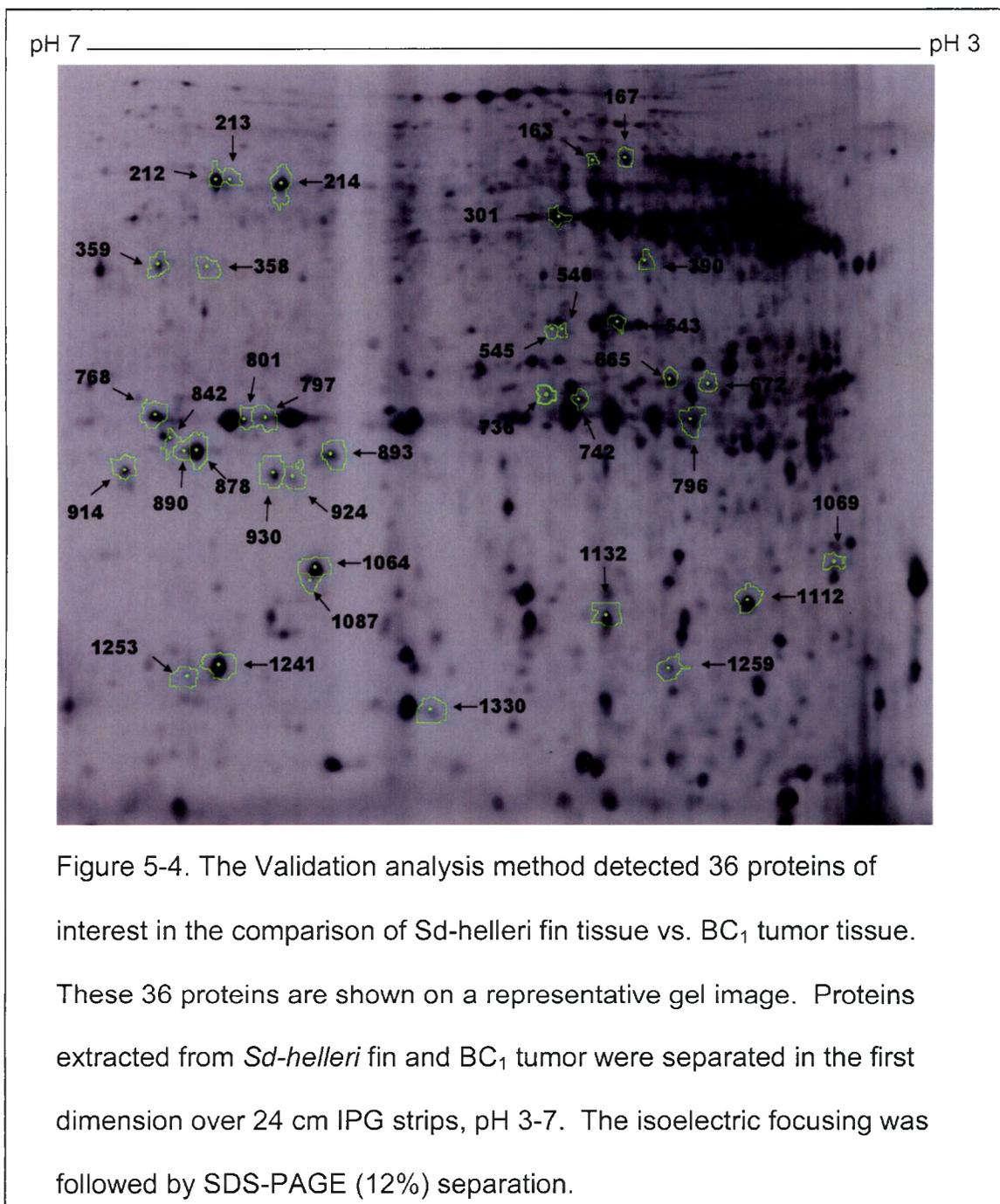


Figure 5-4 is a representative gel image showing the 36 POIs detected using the Validation analysis method. These 36 POIs are spread evenly across

the pH range with some concentration in the middle of the molecular weight range which is similar to the POIs identified using the Visual method.

Table 5-1 lists 12 POIs that were found to be in common between both the Visual and the Validation analysis methods. The 12 POIs are listed according to their identification numbers in figure 5-3 (i.e. Visual analysis method) and the corresponding identification numbers in figure 5-4 detected by the Validation analysis method. Abundance ratios for each of the POIs are also presented for comparison. Seven of the POI pairs found to be common between the Visual and Validation methods have abundance ratios within 0.1 of each others. Four of the POIs from the Visual analysis method have abundance ratios within 0.5 of their corresponding POI's abundance ratio derived from the Validation analysis method. POI 1075 from figure 5-3 and its corresponding POI 1112 from figure 5-4 had the largest difference in abundance ratios between the two analysis methods at 1.52.

Table 5-1. Corresponding identification numbers for common proteins of interest between the Visual and Validation methods presented in figure 5-3 and figure 5-4.				
Figure 5-3 ID # ^a	Abundance Ratio ^b		Abundance Ratio ^b	Figure 5-4 ID # ^c
215	2.84		2.69	212
217	2.84		2.76	214
410	2.17		1.89	359
550	3.10		2.69	546
551	-2.24		-2.50	543
753	2.06		2.15	742
768	4.47		4.40	768
872	2.49		2.43	878
910	2.57		2.55	914
945	2.00		1.97	930
1075	-5.78		-7.30	1112
1170	-2.08		-1.99	1241
^a Protein spot number corresponding to POI number assigned in figure 5-3 resulting from the Visual analysis method ^b Abundance ratios for each protein of interest as seen in BC ₁ Tumor in comparison to <i>Sd-helleri</i> ^c Protein spot number corresponding to POI number assigned in figure 5-4 resulting from the Validation analysis method				

Table 5-2 lists 36 POIs detected using the Validation method. The 36 POIs have identification numbers corresponding to the numbers assigned in figure 5-4. An abundance ratio for each POI is listed. POIs exhibiting positive abundance ratios are on the left-hand side and POIs exhibiting negative abundance ratios are on the right-hand side. The abundance ratios represent protein abundance in BC₁ tumor tissue compared to *Sd-helleri* F₁ fin tissue, thus, positive values indicate POIs that are potentially “up-regulated” in the tumor sample relative to F₁ fin proteins. The largest positive ratio was 4.40 for spot 768. Sixteen POIs had positive abundance ratios over 2.0. The largest negative ratio was -7.20 for spot 1112. Eight POIs had larger negative abundance ratios than -2.0 (i.e. -3.72, -7.20).

Table 5-2. Proteins of interest detected using the Validation analysis method that are identified in figure 5-4 .			
Figure 3 ID #	Abundance Ratio	Figure 3 ID #	Abundance Ratio
768	4.40	842	-1.81
1132	3.28	390	-1.89
1087	2.95	736	-1.93
214	2.76	796	-1.95
1064	2.70	1241	-1.99
212	2.69	167	-2.05
546	2.69	1330	-2.28
914	2.55	163	-2.28
545	2.50	665	-2.47
213	2.47	543	-2.50
878	2.43	672	-2.53
1069	2.40	1259	-3.72
797	2.37	1112	-7.30
1253	2.32	^a Protein spot number corresponding to POI number assigned in figure 5-4 ^b Abundance ratios each protein of interest as seen in BC ₁ tumor in comparison to <i>Sd-helleri</i>	
742	2.15		
358	2.11		
801	1.97		
930	1.97		
924	1.97		
890	1.92		
359	1.89		
301	1.89		
893	1.88		

3) Protein Identifications of proteins of interest from the Visual analysis

The Validation method was not completed until after POIs were isolated using the Visual method. Therefore, the identification of the POIs detected by the Validation method is in progress.

Table 5-3 represents the 32 POIs determined by the Visual analysis method by their identification numbers, their abundance ratios of BC₁ tumor in comparison to *Sd-helleri* F₁ fin tissues, the mass of parent ions that were fragmented using MALDI-TOF MS, the amino acid sequences obtained from

fragmentation, and the possible protein identifications obtained from searching the peptide amino acid sequences against public protein databases.

We were unable to obtain amino acid sequences for 8 of the 32 POIs. Amino acid sequences were obtained for the remaining 24 POIs with varying confidence for identification. Fifteen POIs have expect values of less than 1. Nine POIs have expect values of less than 10 but greater than 1. The proteins with the best identification scores were peroxiredoxin ($E = 2 \times 10^{-05}$), enolase ($E = 6 \times 10^{-06}$), glyceraldehyde-3-phosphate dehydrogenase ($E = 6 \times 10^{-06}$), and actin ($E = 1 \times 10^{-06}$).

D. DISCUSSION

DIGE comparison of *Sd-helleri* F₁ interspecies hybrid fin proteins versus BC₁ melanoma tumor fin proteins gives us the opportunity to explore protein abundance changes that may occur due to development of melanoma. Comparison of *X. maculatus* to *Sd-helleri* (described in chapter 4) identified proteins that were differentially abundant due to interspecies hybridization. Some of these same proteins were identified in this analysis of *Sd-helleri* versus BC₁ tumors and it is assumed these common proteins are differentially abundant due to the interspecies hybridization rather than due to the progression of pigmented cells to melanoma. Common proteins were GST, peroxiredoxin, and actin. In contrast, proteins that are differentially abundant only in the *Sd-helleri* F₁ fin versus BC₁ melanoma tumor tissue may be considered potential biomarkers for melanogenesis.

Visual method results (provided in chapter 4) of *X. maculatus* parental fin tissue versus *Sd-helleri* F₁ hybrid fin tissue identified 23 proteins with 11 having positive abundance ratios and 12 having negative abundance ratios. Thus, it is interesting to observe application of the Visual method to the *Sd-helleri* fin vs. BC₁ melanoma fin tissue detected twice as many high abundance proteins than comparison of *X. maculatus* versus *Sd-helleri* (i.e. 24) and a similar number of low abundance proteins (i.e. 8).

Comparison of the Visual method to the Validation method in this chapter indicates the number of POIs detected were very similar (i.e. 32 for Visual and 36 for Validation). The Visual analysis method however, detected three times more high abundance proteins than low, while the Validation method had two times more high than low abundance proteins. Thus, trends of POI numbers and the abundance tendencies were similar and 12 POIs were detected by both methods.

Of the 32 POIs detected using the Visual method, 24 POIs had at least one peptide amino acid sequence that was able to be tentatively identified from comparison with gene or protein public databases. We were not able to obtain peptide sequences for 8 of these 32 POIs.

One of the POIs tentatively identified was Rho dissociation inhibitor protein with an abundance ratio was -2.34 in BC₁ tumor tissue compared to *Sd-helleri* F₁ fin. There are three types of Rho GTPases that regulate cell growth and cell migration (41). Wang *et. al.* (41) has suggested using RhoGAP-Rho chimeras in order to down regulate Rho activity to attempt reversal of growth and

invasive phenotypes of cancer cells. Thus, Rho dissociation inhibitor appears to be a cell signaling protein that may represent a good candidate as a melanoma biomarker in this experimental system.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a protein found to be differentially expressed two times higher in BC₁ melanoma than *Sd-helleri* F₁ interspecies hybrids. GAPDH is involved in carbohydrate metabolism (42). GAPDH is a well-known housekeeping gene whose levels were thought to be unaffected by cancer transformation; however, Stahl *et al.* (38) have shown that GAPDH had two times higher gene expression in invasive melanoma cells compared to noninvasive melanoma. GAPDH would also be an interesting protein to study in different types of tumors within the *Xiphophorus* fish experimental system. GAPDH was not identified as differentially expressed upon interspecies hybridization which makes it a possible biomarker candidate for melanoma. Further investigation of GAPDH genes could tell us if the same is true in invasive versus noninvasive tumors found in *Xiphophorus*.

Enolase is a neuron-specific serum protein that has been used in previous studies to gauge the severity of neuroblastomas in humans (43). We observed three times the amount of enolase in BC₁ melanoma than in *Sd-helleri* F₁ fin tissue. An increase in enolase has also been suggested a possible diagnostic biomarker for acute brain infarctions (44). Thus, increase in enolase abundance in *Xiphophorus* melanoma may be consistent with neural cell reorganization and indicates a potential area of cellular stress.

This proteomic comparison of BC₁ melanoma tumor versus *Sd-helleri* F₁ interspecies hybrid dorsal fin tissue has provided several potential melanoma biomarkers that will be very useful in future studies of induced neoplasia using the *Xiphophorus* genetic system. Rho dissociation inhibitor controls the cell's ability to migrate and grow and could provide us with a marker for cell cycle signaling. GAPDH, a housekeeping gene, should not be overlooked as a possible marker to decipher invasive and noninvasive transformation in melanoma progression. Enolase has been used as a biomarker to detect cell transformation in neuroblastomas and irregularities due to brain infarctions. With additional protein identifications in progress from the POIs selected via the Validation method, our list of potential melanoma biomarkers may only continue to grow.

Table 5-3. Protein identifications of *Sd-helleri* F₁ versus BC₁ tumor DIGE proteins of interest.

Protein of Interest ^a	BC ₁ Tumor Abundance Ratio ^b	SPITC-peptide precursor ion (m/z) ^c	Peptide Sequence ^d	Protein Identification ^e
221	-2.88	1160.292	AVFPS[I/L]VGR	Actin (<i>D. rerio</i> , E = 1.2)
		2005.950	PDGQV[I/L]T[I/L]GNER	Actin (<i>D. rerio</i> , E = .012)
1064	-2.73	1455.656	FF[I/L]SQ[I/L]D	Protein C2orf29 homolog (<i>M. musculus</i> , E = 5.6)
		1737.613	YFNVDCHQ	Laminin beta-2 chain precursor (<i>R. norvegicus</i> , E = 2.1)
		1909.964	H[I/L]EQFND[I/L]SDYCR	Ubiquitin carboxyl-terminal hydrolase (<i>H. sapiens</i> , E = 1.0)
715	-2.34	1503.491	GQVFTA[I/L]VAEVR	Unnamed
		1823.737	ADPTAPNVQVTR	Rho GDP dissociation inhibitor (<i>D. rerio</i> , E = 0.002)
551	-2.24	1230.201	V[I/L]VE[I/L][I/L]SSR	Unnamed
		1931.381	FVT[I/L][I/L]VAR	Lysosome membrane protein (<i>R. norvegicus</i> , E = 1.2)
		2163.923	QWAQEGN[I/L]ETDAQ	Vacuolar transporter chaperone (<i>S. cerevisiae</i> , E = 0.27)
1170	-2.08	1483.462	[I/L]AANVE[I/L]VVHR	Vesicle trafficking protein (<i>C. elegans</i> , E = 3.2)
		2005.547	SYE[I/L]PDGQV[I/L]T[I/L]GNER	Actin (<i>D. rerio</i> , E = 1 x 10 ⁻⁰⁶)
667	-1.98	1695.513	GD(F,G)VNDNA[I/L][I/L]R	Aminomethyltransferase (<i>H. sapiens</i> , 2.7)
		1711.727	V[I/L]DNA[I/L][I/L]R	Aminomethyltransferase (<i>H. sapiens</i> , 2.7)

830	-1.88	1334.222	V[I/L][I/L]HYFDGR	Glutathione S. transferase (<i>O. cuniculus</i> , E = 0.013)
		2028.522	C[I/L]VHEMT[I/L]GGER	Rac prophage (<i>E. coli</i> , E = 2.5)
		2178.424	QAQRNAEQ[I/L]TS[I/L]R	Polyglutamine-containing protein (<i>D. rerio</i> , E = 1.3)
820	1.80	2028.522	C[I/L]VHEMT[I/L]GGER	Rac prophage (<i>E. coli</i> , E = 2.5)
		2178.424	QAQRNAEQ[I/L]TS[I/L]R	Polyglutamine-containing protein (<i>D. rerio</i> , E = 1.3)
554	1.81	1230.251	V(D,P)E[I/L][I/L]SSR	DNA-methyltransferase (<i>C. elegans</i> , E = 2.3)
		1685.418	MTEDTHLQQLQR	Intersectin-1 (<i>X. laevis</i> , E = 0.23)
		1932.373	[I/L]ETDAF[I/L]GNAR	Cytochrome (<i>C. elegans</i> , E = 5.3)
764	2.10	2179.034	SEAQRNAEQ[I/L]TS[I/L]R	Polyglutamin-containing protein (<i>D. rerio</i> , E = 0.97)
410	2.17	1994.300	[I/L][I/L]SWYDNEYGYSNR	Glyceraldehyde-3-phosphate dehydrogenase (<i>M. musculus</i> , E = 6×10^{-06})
791	2.18	2032.573	FPN[I/L]PY[I/L][I/L]DGDR	Glutathione S-transferase M (<i>D. rerio</i> , E = .002)
786	2.34	1402.306	[I/L]S[I/L][I/L]YPATTGR	Peroxiredoxin (<i>D. rerio</i> , E = 0.016)
		1672.958	M[I/L]A[I/L]S[I/L]DSVED	Peroxiredoxin (<i>D. rerio</i> , E = 0.093)
		1817.005	E[I/L]SVQ[I/L]GM[I/L]DPDER	Peroxiredoxin (<i>D. rerio</i> , E = 2×10^{-05})
		2702.614	AFANEAGTP[I/L]PFP[I/L][I/L]ADDQR	Peroxiredoxin (<i>D. rerio</i> , E = 0.53)
811	2.45	1406.451	[I/L]T[I/L]YPATTGR	Peroxiredoxin (<i>D. rerio</i> , E = 0.12)
		1672.368	M[I/L]A[I/L]S[I/L]DSVED	Peroxiredoxin (<i>D. rerio</i> , E = 0.93)
		1816.359	TQ[I/L]GM[I/L]NPDER	Peroxiredoxin (<i>D. rerio</i> , E = 0.093)
		2702.476	ANEAGTP[I/L]PFP[I/L][I/L]ADDQR	Peroxiredoxin (<i>D. rerio</i> , E = 0.038)

215	2.84	1733.556	D[I/L]GVMVSHR	Enolase (<i>D. rerio</i> , E = 1.2)
		2019.492	PSGASTG[I/L]YEA[I/L]E[I/L]R	Enolase (<i>D. rerio</i> , E = 6 x 10 ⁻⁰⁶)
217	2.84	1757.695	DVAAGCVH[I/L]R	Putative protein (<i>C. elegans</i> , E = 1.7)
		2019.770	QSTG[I/L]YEA[I/L]E[I/L]R	Enolase (<i>D. rerio</i> , E = .012)
550	3.10	1230.474	V[I/L]VE[I/L][I/L]SSR	Annexin (<i>C. elegans</i> , E = 1.4)
		2163.403	MEEG[I/L]NETDAQ	Cadherin 89D precursor (<i>D. melanogaster</i> , E = .018)
1122	3.19	1197.639	ST[I/L]HFA[I/L]R	beta-1 3-N-acetylglucosaminyltransferase bGnT (<i>C. elegans</i> , E = 1.6)
		1458.226	YSV[I/L]SFEGDAR	Putative protein (<i>C. elegans</i> , E = 2.5)
		1667.269	FTPAEFVVT[I/L]TAR	Atrial natriuretic peptide receptor B precursor (<i>R. norvegicus</i> , E = 0.85)
768	4.47	1295.279	Q(Q,P)WS[I/L]G[I/L]R	Unnamed
		1704.451	DPAPDMFDR	Similar to myomesin family, member 3 (<i>D. rerio</i> , E = 1.6)
1075	-5.78	1896.492	CDN[I/L]P[I/L]FGFP	Alpha-2,8-sialyltransferase (<i>T. rubripes</i> , E = 6.3)
945	2.00	1279.369	[I/L][I/L][I/L]WSPVSR	Interleukin precursor (<i>H. sapiens</i> , E = 3.0)
		1285.755	[I/L]GDSWEP[I/L]	Interleukin receptor alpha chain precursor (<i>H. sapiens</i> , E = 2.0)
76	2.22	1335.599	D[I/L]VQFVFPR	Solute carrier anion transporter (<i>M. musculus</i> , E = 1.2)

910	2.57	1238.522	VMHPCT[I/L]AR	Solute carrier (<i>M. musculus</i> , E = 2.2)
		1874.720	VYYENVAS[I/L]pSM(Q,V)R	3-dehydroquinase dehydratase (<i>E. faecalis</i> , E = 7.4)
775	1.85	1229.481	TH[I/L]EPYVR	RNA binding protein (<i>H. sapiens</i> , E = 4.9)
Proteins With No Sequence or ID				
1104	1.91			
1305	2.01			
753	2.06			
27	2.11			
1131	2.14			
1124	2.15			
872	2.49			
317	2.63			
^a Protein spot number corresponding to POI number assigned in Figure 2.				
^b Average abundance ratio across 6 analyses representing 2 replicate gels each				
^c Monoisotopic m/z 2.1 TDC matrix used to collect spectra in positive ion mode				
^d Peptide amino acid sequence determined by postsource decay spectra of SPITC-labeled peptides.				
^e Protein identification based on best expect value score from several database searches including NCBI, SwissProt, and FASTA at University of Virginia				

CHAPTER VI

CONCLUSION

This research has presented the development of a validation method for DIGE analysis. Two DIGE comparisons of the dorsal fins of fish involved in the Gordon-Kosswig melanoma model included *X. maculatus* vs. *Sd-helleri* F₁ and *Sd-helleri* F₁ vs. BC₁ melanoma. This study has provided insight into the proteomic effects of interspecies hybridization and the effects of development of melanoma on protein expression.

A validation method for DIGE analysis was developed that improved our ability to select POIs based on statistical significance 80% power. With addition of the power requirement, this validation method takes into account variance in same-protein spot volumes among gel replicates. The DIGE DeCyder analysis accounted for Type I statistical error but lacked a means of accounting for Type II statistical error. The Validation method's additional requirements for detecting statistically significant protein changes fills the gap left by the DeCyder program and adds the consideration for Type II error which further enhances statistical confidence.

Xiphophorus fish involved in the Gordon-Kosswig melanoma model were the focus of this investigation. The first step was to identify proteins that were differentially abundant due solely to the interspecies hybridization between *X.*

maculatus and *X. helleri* (Sarabia) which resulted in *Sd-helleri* F₁ interspecies hybrids. These *Sd-helleri* F₁ interspecies hybrids exhibit phenotypically enhanced pigmentation compared to either of its parents. The pigmentation pattern, *Sd*, is a macromelanophore pigment pattern inherited by the *Sd-helleri* F₁ from *X. maculatus*. In our first set of experiments, we focused on comparing the dorsal fin proteins of the parental, *X. maculatus* to the dorsal fin proteins of the *Sd-helleri* F₁. By comparing the dorsal fin proteins we were able to identify proteins that exhibited altered abundance due to the interspecies hybridization. Examples of proteins altered are transferrin, glutathione-S-transferase, and peroxiredoxin.

Once we identified proteins altered by interspecies hybridization, we examined the dorsal fin protein differences between the *Sd-helleri* fin tissues and the first generation backcross (BC₁) hybrid dorsal fin melanoma tissue. This experiment was aimed at comparing proteins from dorsal fins that expressed enhanced levels of pigmentation (*Sd-helleri*) and dorsal fin cells with greatly enhanced levels of pigmentation as melanomas that eventually led to tissue necrosis of the fin and the surrounding dorsal region (BC₁ melanoma). When proteins of interest were identified, the proteins that were common with the first experimental set were considered to be differentially abundant due to the genetic hybridization process. Proteins that were only observed to be differentially abundant in the second set of experiments between *Sd-helleri* dorsal fin and BC₁ melanoma were considered potential melanoma biomarkers for this experimental

system. Examples of these potential melanoma biomarkers are Rho dissociation inhibitor, glyceraldehyde-3-phosphate dehydrogenase and enolase.

Future plans include the DIGE comparison of fin tissue from *X. helleri* and *Sd-helleri* F₁ hybrids to present a more complete picture of proteins affected by interspecies hybridization. We also plan to complete the identification of the proteins of interest that were detected by the Validation method for both the *X. maculatus* vs. *Sd-helleri* and *Sd-helleri* vs. BC₁ melanoma comparisons. Once those identifications are made we may choose some of the potential melanoma biomarkers and design primers for complete gene sequencing and cloning. These genes can then be studied in expression studies conducted with real-time PCR to compare mRNA levels and examine whether the mRNA levels follow the same abundance trends. Additionally, there are many other *Xiphophorus* species crosses that could be examined by this method.

LITERATURE CITED

- 1) Gordon M. The genetics of a viviparous top-minnow *Platypoecilus*: The inheritance of two kinds of melanophores (1927). *Genetics*, 12, pp253-283.
- 2) Kosswig C. Über Bastarde der Teleostier *Platypoecilus* und *Xiphophorus* (1928). *Z. Indukt. Abstammungs-Vererbungslehre*, 44, pp150-158.
- 3) Walter RB, Sung HM, Obermoeller RD, Mitchell DL, Intano GW, Walter CA. Relative base excision repair in *Xiphophorus* fish tissue extracts (2001). *Mar. Biotechnol.*, 3, pp50-60.
- 4) Mitchell DL, Nairn RS, Johnston DA, Byrom M, Kazianis S, Walter RB. Decreased levels of DNA excision repair in hybrid fish of the genus *Xiphophorus* (2004). *Mut. Res. Photochem. Photobiol.*, 79, pp447-452.
- 5) Setlow RB, Woodhead AD, Grist E. Animal model for ultraviolet radiation-induced melanoma: platyfish-swordtail hybrid (1989). *Proc. Natl. Acad. Sci. USA*, 86, pp8922-8926.
- 6) Setlow RB, Grist E, Thompson K, Woodhead AD. Wavelengths effective in induction of malignant melanoma (1993). *Proc. Natl. Acad. Sci. USA*, 90, pp6666-6670.
- 7) Narin RS, Kazianis S, McEntire BB, Della Colleta L, Walter RB, Morizot DC. A *CDKN2*-like polymorphism in *Xiphophorus* LG V is associated with UV-B-induced melanoma formation in platyfish-swordtail hybrids (1996). *Proc. Natl. Acad. Sci. USA*, 93, pp13042-13047.
- 8) Narin RS, Morizot DC, Kazianis S, Woodhead AD, Setlow RB. Nonmammalian models for sunlight induced carcinogenesis: genetic analysis of melanoma formation in *Xiphophorus* hybrid fish (1996). *Photochem. Photobiol.*, 64, pp440-448.
- 9) Nairn RS, Kazianis S, Della Coletta L, Trono D, Butler AP, Walter RB, Morizot DC. Genetic analysis of susceptibility to spontaneous and UV-induced carcinogenesis in *Xiphophorus* hybrid fish (2001). *Mar. Biotechnol.*, 3, ppS24-S36.

- 10) Walter RB, Kazianis S. *Xiphophorus* interspecies hybrids as genetic models of induced neoplasia (2001). *ILAR J.*, 42, pp299-322.
- 11) Morizot DC, Siciliano MJ. Linkage group V of platyfishes and swordtails of the genus *Xiphophorus* (Poeciliidae): linkage of loci for malate dehydrogenase-2 and esterase-1 and esterase-4 with a gene controlling the severity of hybrid melanomas (1983). *J. Natl. Can. Inst.*, 71, pp809-813.
- 12) Kazianis S, Gutbrod H, Nairn RS, McEntire BB, Della Colletta L, Walter RB, Borowsky RL, Woodhead AD, Setlow RB, Scharl M, Morizot DC. Localization of a *CDKN2* gene in linkage group V of *Xiphophorus* fishes defines it as a candidate for the *DIFF* tumor suppressor (1998). *Genes Chromosom. Cancer*, 22, pp210-220.
- 13) Gutbrod H, Scharl M. Intragenic sex-chromosomal crossovers of *Xmrk* oncogene alleles affect pigment pattern formation and the severity of melanoma in *Xiphophorus* (1999). *Genetics*, 151, pp773-783.
- 14) Scharl M. Homology of melanoma-inducing loci in the genus *Xiphophorus* (1990). *Genetics*, 126, pp1083-1091.
- 15) Ajuha MR, Lepper K, Anders F. Sex chromosome aberrations involving loss and translocation of tumor-inducing loci in *Xiphophorus* (1979). *Experientia*, 35, pp28-30.
- 16) Nanda I, Volf JN, Weis S, Korting C, Froschauer A, Schmid M, Scharl M. Amplification of a long terminal repeat-like element on the Y chromosome of the platyfish, *Xiphophorus maculatus* (2000). *Chromosoma*, 109, pp242-249.
- 17) Wittbrodt J, Adam D, Malitschek B, Maueler W, Raulf F, Telling A, Robertson SM, Scharl M. Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus* (1989). *Nature*, 341, pp415-421.
- 18) Wellbrock C, Fischer P, Scharl M. P13-kinase is involved in mitogenic signaling by the oncogenic receptor tyrosine-kinase *Xiphophorus* melanoma receptor kinase in fish melanoma (1999). *Exp. Cell Res.*, 251, pp340-349.
- 19) Patton WF. Detection technologies in proteome analysis (2002). *J. Chromo. B.*, 771, pp3-31.

- 20) Monteoliva L, Albar JP. Differential proteomics: an overview of gel and non-gel based approaches (2004). Briefings in Functional Genomics and Proteomics, 3, pp220-239.
- 21) Van den Bergh G, Arckens L. Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel-based proteomics (2004). Current Opinion in Biotech., 15, pp38-43.
- 22) Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts (1997). Electrophoresis, 18, pp2071-2077.
- 23) Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I. A novel experimental design for comparative two-dimensional analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard (2003). Proteomics, 3, pp36-44.
- 24) Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology (2001). Proteomics, 1, pp377-396.
- 25) Gustafsson JS, Ceasar R, Glasbey CA, Blomberg A, Rudemo M. Statistical exploration of variation in quantitative two-dimensional gel electrophoresis data (2004). Proteomics, 4, pp3791-3799.
- 26) de Hoffmann E, Stroobant V. Mass Spectrometry: principles and applications (2001). John Wiley and Sons, LTD, New York USA.
- 27) Chen P, Nie S, Mi W, Wang X-C, Liang S-P. *De novo* sequencing of tryptic peptides sulfonated by 4-sulfophenyl isothiocyanate for unambiguous protein identification using post-source decay matrix-assisted laser desorption/ionization mass spectrometry (2004). Rapid Comm. Mass. Spec., 18, pp191-198.
- 28) Oehlers LP, Perez AN, Walter RB. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of 4 sulfophenyl isothiocyanate-derivatized peptides on AnchorChiptm sample supports using the sodium-tolerant matrix 2,4,6-trihydroxyacetophenone and diammonium citrate (2005). Rapid Comm. Mass Spec., 19, pp752-758.
- 29) Karp NA, Lilley KS. Maximising sensitivity for detecting changes in protein expression: experimental design using minimal CyDyes (2005). Proteomics, 5, pp3105-3115.

- 30) Gravetter FJ, Wallnau LB. *Statistics for behavioral sciences* (2000). Wadsworth Thomson Learning, Australia.
- 31) Molloy MP, Brzezinski EE, Hang J, McDowell MT, VanBogelen RA. Overcoming technical variation and biological variation in quantitative proteomics (2003). *Proteomics*, 3, pp1912-1919.
- 32) Dell RB, Holleran S, Ramakrishnan R. Sample size determination (2002). *ILAR J.*, 43, pp207-213.
- 33) Karp NA, Kreil DP, Lilley KS. Determining a significant change in protein expression with DeCyder™ during a pair-wise comparison using two-dimensional difference gel electrophoresis (2004). *Proteomics*, 4, pp1421-1432.
- 34) Garbis S, Lubec G, Fountoulakis M. Limitations of current proteomic technologies (2005). *J. of Chromo. A.*, 1077, pp1-18.
- 35) Gaeta BA. Database searching using short peptide sequences (1997). *ABRF News*.
www.abrf.org/ABRFNews/1997/December1997/dec97tips.html
- 36) Lambert LA, Perry H, Meehan TJ. Evolution of duplications in the transferrin family of proteins (2005). *Comp. Biochem. and Physio. B: Biochem. and Mol. Bio.*, 140, pp11-25.
- 37) Kang JS, Cho D, Kim YI, Hahm E, Kim YS, Jin SN, Kim HN, Kim D, Hur D, Park H, Hwang YI, Lee WJ. Sodium ascorbate (vitamin C) induces apoptosis in melanoma cells via the down-regulation of transferrin receptor dependent iron uptake (2005). *J Cell Physiol.*, 204, pp192-197.
- 38) Stahl S, Bar-Meir E, Friedman E, Regev E, Orenstein A, Winkler E. Genetics in melanoma (2004). *Isr. Med. Assoc. J.*, 6, pp774-777.
- 39) Noguez MR, Giralt M, Cervello I, Del Castillo D, Espeso O, Argany N, Aliaga A, Mallol J. Parameters related to oxygen free radicals in human skin: a study comparing healthy epidermis and skin cancer tissue (2002). *J. Invest. Dermatol.*, 119, pp645-652.
- 40) Groten K, Dutilleul C, van Heerden PDR, Vanacker H, Bernard S, Finkemeier I, Dietz K-J, Foyer CH. Redox regulation of peroxiredoxin and proteinases by ascorbate and thiols during pea root nodule senescence (2006). *FEBS Letters*, 580, pp1269-1276.

- 41) Wang L, Yang L, Lou Y, Zheng Y. A novel strategy for specifically down-regulating individual Rho GTPase activity in tumor cells (2003). *J. Bio. Chem.*, 278, pp44617-33625.
- 42) Bosworth CA, Chou C-W, Cole RB, Rees BB. Protein expression patterns in zebrafish skeletal muscle: initial characterization and the effect of hypoxic exposure (2005). *Proteomics*, 5, pp1362-1371.
- 43) Sandoval JA, Dobrolecki LE, Huang J, Grosfeld JL, Hickey RJ, Malkas LH. Neuroblastoma detection using serum proteomic profiling: a novel mining technique for cancer? (2006). *J. Pediatric Surgery*, 41, pp639-646.
- 44) Selakovic V, Raicevic R, Radenovic L. The increase of neuron-specific enolase in cerebrospinal fluid and plasma as a marker of neuronal damage in patients with acute brain infarction (2005). *J. Clinical Neuroscience*, 12, pp542-527.

VITA

Amy Nichole Perez was born in Mesquite, Texas, on September 21, 1981, the daughter of Julie Ann Perez and David Joseph Perez. After completing her work at Lancaster High School, Lancaster, Texas, in 2000, she entered Texas State University–San Marcos. She received a Bachelor of Science from Texas State University–San Marcos in 2004. In January 2004, she entered the Graduate College of Texas State University–San Marcos.

Permanent Address : 202 Pebblecreek
 Red Oak, Texas 75154

This thesis was typed by Amy Nichole Perez.