RETICULITERMES FLAVIPES

(ISOPTERA: RHINOTERMITIDAE)

PROTEOMICS

By

CHARLES JERRY BOWEN, JR.

Bachelor of Science
Midwestern State University
Wichita Falls, Texas
1994

Master of Science
Midwestern State University
Wichita Falls, Texas
2002

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May 2007

RETICULITERMES FLAVIPES

(ISOPTERA: RHINOTERMITIDAE)

PROTEOMICS

Dissertation Approved:

Bradford M. Kard

Dissertation Advisor

Jack W. Dillwith

Kristopher L. Giles

Mark E. Payton

A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Jack W. Dillwith for all he has done to help me with this project. Without his guidance, this project would have been impossible to complete. Dr. Brad Kard was invaluable as my committee chair and major advisor. Also, I would like to extend my sincerest gratitude to Dr. Kris Giles and Dr. Mark Payton for their participation on my committee.

I would like to thank Robin Madden and Dr. Bryna Donnelly for both their technical assistance 'at the bench' and their friendship.

I would like to thank the people of the Department of Entomology and Plant Pathology for allowing me to be a part of the whole and their patience in waiting to apply a termiticide until my research was complete. My thanks to the department head, initially, Dr. Russell Wright during the early days, and later, Dr. Jonathan Edelson for their support.

I would like to thank Oklahoma State University and The Nature Conservancy for allowing me to collect termites on their facilities. Thank you to Michael P. Doss and Greg H. Broussard for assistance collecting 'wild' termites from the field. Also, thank you to Doug Keuhl for assistance sorting termite instars in the laboratory.

I would like to thank Don Arnold for his patience when the termites I was collecting damaged several of the slide boxes.

I would like to thank Dr. Steve Hartson, Janet Rogers, and Lisa Whitworth of the Oklahoma State University Biochemistry and Molecular Biology Core Facility. I would also like to thank Dr. Patricia Ayoubi and Dr. Peter Hoyt for their assistance assimilating the molecular functions. I would also like to thank Hua Weng for maintaining and updating local protein databases.

I would like to thank my grandmother, Clara Riordan, my parents, Don and Joan Riddles, and Charles and Elizabeth Bowen. I would also like to thank Pam Hathorn and my friends for their support over the past several years. Without you, I could never have made it this far.

I would like to thank Matrix Science, Inc. (Mascot) and Syngene Ltd.(Dymension) for their consent to reproduce images generated by their software.

To those whom I have failed to name, I thank you as well.

TABLE OF CONTENTS

Chapter **Page**

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

ABC	Ammonium bicarbonate
ACN	Acetonitrile
$AgNO_3$	Silver nitrate stain
BLAST	Basic Local Alignment Search Tool
BMB	Biochemistry and Molecular Biology
CAF	Chemically assisted fragmentation
CBB	Coomassie brilliant blue dye
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHCA	α-cyano-4-hydroxycinnamic acid
Da	Dalton(s)
ddH_2O	Water filtered by reverse osmosis, de-ionized, the purified by distillation
DIGE	Differential in-gel electrophoresis
dpi	Dots per inch
DTT	Dithiothreitol
EST	Expressed sequence tag
HTML	Hyper-Text Markup Language
IEF	Isoelectric focusing
IPG	Immobiline pH Gradient
MALDI	Matrix assisted laser desorption/ionization
MOWSE	Molecular weight search
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
M_w	Molecular weight
m/z	Mass-to-charge ratio
NCBI	National Center for Biotechnology Information
NL	Non-linear
PAGE	Polyacrylamide gel electrophoresis
PDA	Piperazine diacrylamide
pI	Isoelectric point

PMF	Peptide mass fingerprint
SDS	Sodium dodecyl sulfate
TCA	Trichloroacetic acid
TEMED	N,N,N^{1},N^{1}-tetramethylethylenediamine
TOF	Time-of-flight

INTRODUCTION

Termites cause 2 to 3 billion dollars annually in damage to wooden buildings and structures in the United States of America (1998-2004). In Oklahoma, termites cause several million dollars in damage to wooden structures each year (Criswell et al. 2001). Accordingly, prevalent areas of termite research relate to management of these destructive insects. Effectiveness of control measures is typically evaluated through field testing using in-ground and soil-surface monitoring stations. Field testing is expensive and time consuming, and investigations of several termiticides including soil-applied and above-ground insecticides, baits incorporating insect growth regulators, and wood preservatives are on-going. Protein analysis of subterranean termites has received little attention, but may eventually yield biomarkers that would provide a valuable method of evaluating the effectiveness of termiticides (Veenstra et al. 2005, Veenstra 2007). A biomarker is any metabolite that undergoes change(s) when exposed to various environmental effects (Redei 2003). Biomarkers are identified through the comparison of metabolite samples under different conditions. For example, medical research uses proteomics technology to compare healthy with diseased tissue, providing identification of biomarkers for a particular condition (Bukowska et al. 2004, Waldburg et al. 2004, Immler et al. 2006). This technology could potentially be applied to identify biomarkers in termites exposed to a termiticide, or to compare effects of a termiticide on specific protein(s) found between termites not exposed to a termiticide. This technology could

lead to safer termiticides, having a very specific target, reducing off-target effects. Another potential benefit of proteomics could be the identification of a termite protein(s) specific to particular species of termites. If such proteins can be identified, it would be possible to use proteomics as a taxonomic tool.

The most significant contribution of termite proteome research would be to help elucidate the fundamental biology of termites. Knowledge provided by fundamental research enhances our ability to understand the workings of applied research. The strength of proteomic studies ranges from the ability to explore the presence of proteins within a single cell to mapping the proteins found throughout an entire organism(Pennington et al. 1997). For example, proteomics techniques are used to study proteins considered essential for the survival of liver flukes within a host (Boonmee et al. 2003), and to study cellular effects caused by exposure to various concentrations of metals (Hu et al. 2003), as well as to evaluate protein targets for drugs (Archakov et al. 2003).

In the human health arena, GeneProt (Geneva, Switzerland) is a laboratory conducting proteomic research comparing healthy with unhealthy human tissues. The differences between these tissues may hold the key to finding a cure for some diseases (Service 2001). In humans, many drugs affect proteins (Palzkill 2002). Therefore, proteomic techniques are being used both to identify biomarkers and to validate drug effectiveness (Bodovitz and Joos 2004). Differential identification between tissue samples containing tumor tissue and normal tissue is now possible in some cases (Celis et al. 1999, Adam et al. 2003, Dwek and Alaiya 2003).

It is probable that many currently used termiticides detrimentally affect insect proteins. The ability to identify these detrimental effects on specific internal termite proteins may provide insight into the potential of specific termiticides as termite management tools. Isolation and identification of a single protein or several proteins that are affected by a specific chemical would increase the possibility of identifying chemicals that target specific insects. Without a baseline protein profile, differential studies are impossible. The establishment of a standardized protein map would allow comparison among various studies based on this technology. While proteomics has extensive potential, it is impossible to accurately predict how fundamental research may impact future research.

Liebler (2002) defines proteomics as "the study of multi-protein systems". Proteomics has been used to elucidate the proteins occurring in an organism exposed to specific conditions. This technology has also been used to compare proteins expressed in healthy tissues with proteins expressed in diseased tissues. Examples of using proteomics in a clinical research setting are to identify biomarkers in cancer (Petricoin and Liotta 2004, Bertucci et al. 2006), lung disease (Waldburg et al. 2004), heart disease (Bukowska et al. 2004), and neurological disease (Nielsen et al. 2005) research. Differential proteomics allows scientists to compare the changes and explore responses in cells exposed to different stimuli and explore how a cell responds to those stimuli (Kahn 1995). Besides clinical uses, proteomics can be used for fundamental scientific research. Cash (1998) used proteomics to characterize bacterial proteomes, while Giavalisco et al. (2005) observed various tissues found in *Arabidopsis thaliana*

(Thale Cress). Proteomics has also been used to investigate the protein composition of *Apis mellifera* royal jelly (Scarselli et al. 2005).

Proteomics facilitates the elucidation of a group of proteins rather than a single protein. In a living organism, a protein profile is dynamic and can change rapidly, affected by either internal and/or external environmental influences. Therefore, stable laboratory conditions are an important factor in minimizing profile variations during experiments. The effect of termiticidal compounds on termite internal proteins has not been investigated, and a baseline protein profile for *Reticulitermes flavipes* (Kollar) has not been determined. Thus, comparisons of proteins in *R. flavipes* exposed to termiticides with proteins in non-exposed termites cannot be done at this time. Once a basic protein profile is determined for *R. flavipes*, protein analyses may provide an evaluation method for efficacy of termiticides.

Two-dimensional (2-D) electrophoresis, first described by O'Farrell (1975), is a primary technique used for protein elucidation. Potentially, 2-D electrophoresis can discriminate between similar protein molecules that possess a difference of a single electrical charge (O'Farrell 1975). Thus, this technique has the capability to produce extensive amounts of raw information (Vihinen 2003). After conducting 2-D electrophoresis, the resulting protein spots in the 2-D gel are excised, and in-gel digestion procedures are conducted. The resulting peptides are extracted from the gel and prepared for mass spectrometry (James et al. 1993). The peptides are mixed with recrystallized alpha-cyano-4-hydroxycinnamic acid (CHCA) matrix and placed on a stainless-steel plate that is then inserted into a matrix-assisted laser desorption/ionization – time-of-flight mass spectrometer (MALDI-TOF MS). This instrument is frequently used for

measuring the mass-to-charge ratio (*m/z*) of ions. A laser is used to desorb (vaporize) and ionize the peptide. Ions having larger *m/z* values travel faster through the column vacuum than ions with smaller *m/z* values. This effect separates the ions by their *m/z*. As these ions strike the detector, a list of peaks is generated for the protein that is termed a peptide mass fingerprint (PMF). PMFs are specific for peptides (Simpson 2003) and can be used to identify proteins by searching protein databases (Liebler 2002).

Several laboratory methods must be used to accurately and clearly map proteins. The primary goal of this research is to establish a system that provides consistent and reproducible results for use in the long-term study of termite protein profiles. Additionally, establishment of an *R. flavipes* protein database will allow protein comparisons among termite species, and among different colonies of the same species, as well as during 'time-point' experiments within the same colony.

OBJECTIVES

I. Develop a system for analyzing the termite proteome

II. Develop standard reference protein map for *Reticulitermes flavipes*

III. Initial characterization of the putative *R. flavipes* proteome

IV. Test for differential protein expression among *R. flavipes* colonies

V. Test for differential protein expression between worker and soldier castes

LITERATURE REVIEW

Reticulitermes flavipes (Kollar) is a paurometabolous insect located taxonomically in the family Rhinotermitidae, Isoptera. Isoptera means "equal wing" and is derived from the reproductively mature termite characteristic of having two pair of approximately symmetric wings (Krishna 1989). The accepted common name for *Reticulitermes flavipes* (Kollar) is the "eastern subterranean termite" (ESA 2004). *R. flavipes* is found throughout most of the continental United States (Snyder 1954, Gold et al. 1999).

Termites are economically important due to their wood-destroying feeding habits. In nature, the eastern subterranean termite forages for cellulose-containing material, such as wood, and it is common for termites to forage into man-made structures. As a result, they destroy structural components, books, paper, and many other products (Spear 1970). Their ability to digest cellulose is facilitated by a mutualistic symbiotic relationship with a variety of microorganisms from the Archaea, Eubacteria, and Eucarya domains (Honigberg 1970, Bignell 2000, Inoue et al. 2000).

Termites are eusocial, living cooperatively in colonies. Colony structure is organized by castes, or physical forms specialized for different biological functions. Some other eusocial insects, for example, ants and bees, have haplodiploid castes comprised of a single sex. In termites, both sexes are diploid (Thorne and Traniello 2003) and each caste contains both males and females (Bennett et al. 1988, Krishna 1989, Laine and Wright 2003). Most termite species have a complex life cycle and accurate

definitions of life cycle stages remain under debate (Laine and Wright 2003).

As an example, a simplified termite life cycle consisting of egg, larva, nymph, worker, soldier, and reproductive will be described. Each year when seasonal conditions are favorable, swarms of winged reproductives, or alates, emerge from the colony. A male and female alate form a tandem pair and shed their wings via a suture near the base of the wings. This pair becomes a founding queen and king that select a nesting site to begin a new colony and initially to feed and care for their first brood.

Termites have an uncommon ability to molt into different castes. The first two instars after eggs hatch are called larvae. Caste differentiation is thought to begin at the third instar (Laine and Wright 2003), although it may begin in the egg (Snyder 1925, Miller 1969). Most of the early young will develop as workers (Krishna 1969). Some workers progress to become soldiers or reproductives (Laine and Wright 2003). Workers forage for cellulose sources, build nesting and protective structures, and feed the soldier and reproductive termites via trophallaxis. Over time, a few workers undergo morphological changes to form a presoldier (callow soldier), and then a soldier (Miura 2001). Soldiers are highly specialized for colony defense, possessing an enlarged head capsule bearing large and powerful mandibles. Soldiers defend against enemies such as ants or other outside colony foraging termites (Bennett et al. 1988, Krishna 1989).

In Oklahoma, subterranean termites are distributed statewide, with *Reticulitermes* being the primary indigenous genus, although termites from other genera have occasionally been collected (Snyder 1954, Weesner 1970, Nutting 1990, Brown et al. 2004). The most current Oklahoma survey (Brown et al. 2004) confirmed four indigenous species: *R. flavipes, R. hageni* , *R. tibialis*, and *R. virginicus*. Species

identification is most definitive when both alates and soldiers are examined. However, when alates are not available, soldiers alone can be identified. Determinations of genetic markers for species identification are well underway (Austin et al. 2004).

Proteins are polymers constructed from long sequences of monomers called amino acids. Each amino acid has the basic structure of HOOC CH_2 NH_3^+. This structure has an amide (NH_3) group on one end and a carboxylic acid (COOH) on the opposing end. The difference among amino acids is the side chain structure attached at the second carbon. There are approximately twenty amino acids formed by different side chains. These may be as simple as a single hydrogen atom as in glycine, more complex such as the methylbenzene in phenylalanine, or even additional amine (NH_3) groups.

On a cellular level, proteins are synthesized in the cytoplasm by ribosomes, cellular structures which link the amino acids together, forming chains. These are connected by peptide bonds between the carboxylic acid of one amino acid and the amide on the next amino acid. Short chains of amino acids are referred to as peptides. Peptides are combined to form the primary protein structure. The structure and function of a protein depend on its amino acid sequence. Protein structure consists of four basic levels: primary, secondary, tertiary, and quaternary. Hydrogen-bonding causes the chains to form secondary structures, i.e., the α-helix, the β-sheet, and turns (Petsko and Ringe 2004). These secondary structures are combined to form folded tertiary structures (Boyer 2002). When multiple tertiary structures are combined, they form quaternary structures (Petsko and Ringe 2004).

As the peptide chains bend and overlap, they link to create a semi-rigid structure, called a conformation. Each different protein conformation allows the protein to interact

with other molecules in a 'lock and key' style arrangement. This arrangement limits protein interactions to specific molecules. Proteins serve a variety of functions that include controlling gene expression, use as a molecular-level identification system by many immune systems, and serving as structural molecules for tissues such as muscle (Finkelstein and Ptitsyn 2002). The most important function of proteins may be their ability to function as enzymes that catalyze chemical reactions. Each cell may contain thousands of proteins that fulfill enzymatic functions (Mathews et al. 2000).

Typically, traditional protein biochemistry investigates characteristics of a single protein, whereas proteomics examines proteins on a comprehensive basis (Pandey and Mann 2000). Levels of protein expression are constantly changing within a cell or organism (Corthals et al. 2000). Thus, proteomics is the study of the protein complement of an organism's genome at a specific moment in time. Therefore, proteomics is an aspect of 'functional genomics'. Proteomics investigates molecular characteristics of an organism that are not evident using traditional molecular biology techniques (Gygi et al. 1999a). Proteomics can visualize the up-regulation or down-regulation of protein expression as affected by external influences (Zivy and de Vienne 2000) as well as protein-protein interactions, protein expression, and protein modifications (Gygi et al. 1999a, Palzkill 2002, Graves and Haystead 2003). However, protein expression levels cannot be predicted by the genome (Gygi et al. 1999a).

Proteomic techniques combine two-dimensional (2-D) electrophoresis, image analysis, mass spectrometry, and bioinformatics to visualize and analyze a protein profile. While there are limitations, 2-D gels currently remain the most powerful method for resolving complex mixtures of proteins. Examples of these limitations of the 2-D

electrophoresis system are its minimal ability to visualize proteins at extreme pH ranges (Giorgianni et al. 2003), possible masking of sparse proteins by more abundant proteins (Gygi et al. 1999b), and the difficulty of dissolving hydrophobic molecules (Pandey and Mann 2000). The potential of proteomics was demonstrated through early studies. For example, Klose and Kobalz (1995) used proteomic techniques for genomic analysis, Posch et al. (1995) used proteomics to analyze the genetic variability of carrot seed proteins, and Shevchenko et al. (1996) used proteomic techniques to investigate silver staining protocol modifications. However, proteomics is maturing into an important technology for analyzing complex samples such as bodily fluids or tissues (Domon and Aebersold 2006).

Two-dimensional electrophoresis has the potential to differentiate protein molecules with a difference of a single charge. This is accomplished by the separation of denatured proteins by their isoelectric point (pI) in the first dimension, and, by their molecular weight (M_w) in the second dimension (O'Farrell 1975). Isoelectric point separation in solution was described by Kolin (1955). Later, O'Farrell (1975) combined isoelectric focusing with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and also described 2-D electrophoresis. Klose (1999) developed a 40cm-wide 2-D gel electrophoresis system capable of resolving over 10,000 proteins in a single separation.

Early methods of isoelectric separation were undertaken using 2.0% carrier ampholines in gels cast in glass tubing (~2.5mm inside diameter), or strips of gel cut from a slab gel (Ferro-Luzzi Ames and Nikaido 1976). Traditionally, these tube gels, (polyacrylamide gels cast in tubes), were used to generate pH gradients. However, these

pH gradients were not stable over time and suffered the additional effect of being easily damaged during handling. These methods produced variable results that were not easily reproducible. Currently, gel strips having immobilized pH gradients are commonly used. The use of commercially available immobilized pH Gradient (IPG) strips have increased reproducibility in the first dimension (Giorgianni et al. 2003). IPG strips consist of an acrylamide gel (infused with acrylamido buffers) coated on a plastic backing and cut into 3mm-wide strips (Anonymous 1998, Fichmann 1999). The plastic backing strengthens the gel for handling. These strips are dehydrated and frozen prior to packaging. A buffer gradient is integrated into the gel during casting. Unlike tube gels, IPG strips are stable over time. However, temperature fluctuations and exposure to CO_2 can detrimentally affect the immobilization (Amersham Biosciences 1998)

IPG strips are rehydrated (reswelled) to the original gel thickness before conducting isoelectric focusing using passive rehydration methods described by Gorg and Weiss (1999). Typically, proteins are infused into the IPG strip during reswelling. IPG strips can also be reswelled using active rehydration, or by applying voltage through the strip during rehydration. The protein sample is combined with the reswelling solution and both are absorbed into the strip as it rehydrates. Rehydrated IPG strips are loaded into the focusing unit and voltage is applied to create a uniform electrical field. The electrical field has a voltage gradient where one end is positively charged and the other is negatively charged. Proteins are amphoteric, thus, their net electrical charges change dependent upon the pH of the surrounding environment. The voltage gradient causes the proteins to migrate along the IPG strip until the proteins stops migrating at the pH where their net charge is zero (Berkelman and Stenstedt 1998). This is the isoelectric point (pI)

or where the difference between the positive and negative charged amino acids (residues) is zero (Bruening et al. 1970). When the protein is no longer charged, it ceases migrating along the gradient. IPG strips which have been exposed to the electrical charge until the proteins have migrated until reaching a zero charge are termed 'focused'. When external current is removed from a pH gradient, the proteins begin to drift away from their pI (Amersham Biosciences 1998).

In the second dimension, sodium dodecyl sulfate (SDS) is an anionic detergent used to equilibrate 'focused' IPG strips. SDS breaks non-covalent bonds, and bonds to the surface of the protein molecules to create a uniform charge-to-mass ratio for each protein. The uniform charge causes the protein molecules to move through the acrylamide gel based on M_W rather than electrical charge (Redei 2003). Thus, the second dimension is completed by the additional separation of focused protein molecules by M_W.

This is accomplished by first placing the edge of the focused IPG strip in contact with the top edge of the SDS polyacrylamide gel so that edges have maximum contact. Electrical voltage is applied to the electrophoresis unit and electrons move through wire to buffer tanks. In the buffer tanks, the electrons are conducted through the solution by ionic movement. The gel components include a buffer, allowing electron movement within the gel. This establishes a voltage gradient across the gel and facilitates the movement of proteins (Bruening et al. 1970).

Protein separation is facilitated by the pore size in the gel. Pore size is a function of the acrylamide concentration (Bruening et al. 1970) and crosslinker concentration (Kabiri et al. 2003). The acrylamide molecules form long chains, bonded together with cross-links. Two common crosslinkers used in conjunction with acrylamide are

bis-acrylamide and piperazine diacrylamide (PDA). The crosslinkers function to form crosslinks between the acrylamide chains. As the crosslinker concentration increases, it can also decrease gelation time (Kabiri et al. 2003). As the acrylamide concentration increases, the cross-links also become more abundant. As the number of cross-links increases, the pore size diminishes. Smaller proteins move through the pores quickly, while large proteins require more time, eliciting separation to occur (Bruening et al. 1970).

After the second dimension is completed, some method of visualization must be used to elucidate proteins in the gel. Staining techniques are used to visualize proteins. Two of the most common non-fluorescent positive stains used for protein visualization are silver nitrate and Coomassie brilliant blue (Yan et al. 2000). Silver nitrate staining is approximately 100 times more sensitive than Coomassie staining (Rabilloud 1999). There are also negative staining methods, for example, Imidazole-zinc, which stains the background rather than the protein (Matsui et al. 1999).

After staining, the gel is transferred to a flatbed scanner connected to a personal computer (PC) and an image is recorded as a grayscale uncompressed tag(ged) image file format (TIFF) file. Several software image analysis programs, such as ImageMaster™, HT Analyzer™, PDQuest™, or Dymension™, can be used to analyze the digital image. Spot detection is accomplished by applying mathematical algorithms, such as Gaussian modeling that assigns numbers to the grayscale density of the pixels of the image (Appel and Hochstrasser 1999). Dymension (Syngene, Frederick, MD, USA) software was used to analyze gels at the Oklahoma State University (OSU) Insect Biochemistry Lab. This software can be used for determining the location and estimate density of the protein spot

in the gel, and for alignment of multiple gels for comparative analysis. The software was also used to catalog the coordinates of each spot using a Cartesian coordinate system. In this system the x-axis corresponds to the pI and the y-axis corresponds to the molecular weight. This aids in locating corresponding spots on different gels.

Several options are available for protein identification, including traditional biochemistry methods or newer methods involving mass spectrometry. At OSU, mass spectrometry is used to yield putative protein identifications. The mass spectrometry is undertaken using a Matrix Assisted Laser Desorption/Ionization-Time of Flight (MALDI-TOF) mass spectrometer. Protein spots are excised from the gel and the proteins are digested to peptides in the gel prior to extraction. After the peptides are extracted, a peptide mass fingerprint (PMF) is generated using the MALDI-TOF instrument.

The PMFs acquired during mass spectrometry are compared with peptide databases containing identified proteins (Courchesne and Patterson 1999, Wilkins et al. 1999) using software to search various databases, comparing peak data. The probability of a protein being correctly identified is based on scoring. Different database search engines use different scoring algorithms to determine the most likely protein identification based on the peptide masses submitted (Gras et al. 2003). Protein identification from peptide databases is still somewhat provisional, and is not as irrefutable as are traditional biochemistry methods.

Another mass spectrometry technique for analyzing protein is post-source decay (PSD). PSD refers to a collection of several methods: sample activation, in-source activation, and post-source activation. These methods allow for peptide sequencing without requiring a chromatographic separation of the protein sample (Spengler 2001).

Basically, a protein is digested with trypsin, and a specific peptide is selected and separated from the rest of the peptides using the ion gate of the MALDI-TOF instrument. Further fragmentation of the selected peptide produces a spectrum that allows the elucidation of its amino acid sequence. The amino acid sequence and the molecular mass of the peptide fragment are used to identify the protein using existing protein databases (Marekov and Steinert 2003).

By using groups of peptides to facilitate, mass spectrometry can be conducted on the N-terminus or the C-terminus of the peptide (Marekov and Steinert 2003). For example, sulfonation of peptides is used to enhance their fragmentation of the peptides. Chen et al. (2004) discuss a method of sulfonation that 'tags' the peptide by attaching a negative sulfonic group to the N-terminus of the peptide. This tagging facilitates the timely decay (fragmentation) of the peptide into amino acids, thus improving the accuracy of the PSD technique (Chen et al. 2004).

Recent advances in proteomic techniques include isotope-coded affinity tagging (ICAT) and differential in-gel electrophoresis (DIGE). ICAT combines a protein sample labeled by a stable isotope with a non-labeled protein sample and then separates these proteins using 2-D electrophoresis. The resulting individual protein samples in a given spot can then be identified by mass spectrometry. ICAT is also useful for quantifying protein abundance (Yu et al. 2004). DIGE uses fluorescent markers to compare multiple protein samples within the same gel and eliminates spot alignment problems (Alban et al. 2003, Hu et al. 2003, Knowles et al. 2003, Yan and Marriott 2003).

Another popular recent proteomic technique is the use of microarray technology. Microarrays are suited for research requiring high-throughput of samples. One nanoliter

volumes of solutions containing various protein samples are distributed on a glass microscope slide in an array pattern and referred to as a 'chip' (Hoy 2003). This allows thousands of samples to be processed simultaneously. Typically, a fluorescent dye is added to each sample. As proteins are expressed, the dye becomes more intense. The dye intensity is measured as an indicator of expression (Churchill 2002). Microarrays are frequently used to study mRNA interactions, however, protein chips are being developed (Brooksbank 2000).

Another useful tool for elucidating protein presence is ^{18}O labeling. This technique is applied during the hydrolysis step of protein digestion (Stewart et al. 2003). Separate protein samples are digested in ^{16}O or ^{18}O water buffer. These digestions are then mixed so that half of each isotope-labeled sample is added to half of the non-labeled sample. This allows each protein sample to act as an internal standard for the other sample (Wang et al. 2001). Mass spectrometry is used to differentiate between labeled and non-labeled proteins (Yao et al. 2001). This technique is useful for identifying cross-linked protein residues in studies of the protein structure (Back et al. 2002) and facilitating glycoprotein identification (Kuster and Mann 1999).

MATERIALS AND METHODS

Materials: All reagents are electrophoresis grade or higher.

Termite collection and preparation – Termites were collected from three geographically separate colonies of *Reticulitermes flavipes*. Colony 1 was located at the Noble Research Center at Oklahoma State University, immediately adjacent to, and interior to the north wall of the K. C. Emerson Entomology Museum. Colonies 2 and 3 were collected on the Nature Conservancy's Tallgrass Prairie Preserve 24.1 km (15 mi) north of Pawhuska, Osage Co., OK.

Groups of termites from each colony were brought into the laboratory and placed into 38-liter steel containers and provisioned with pine (*Pinus* sp.) sap wood (Kard et al. 2003). Termites were acclimated to an environment of approximately 98% relative humidity, 25°C ambient temperature, and no photoperiod. Prior to use, termite groups were maintained under these laboratory conditions for a minimum of 30 days to allow for stabilization and to minimize environmental differences between groups. Termites were harvested as needed.

Objective I – Develop a system for analyzing the termite proteome.

Part 1: Optimize solubilization and staining

Whole-body protein extraction without protease inhibitor –A stock solution was prepared by mixing 0.07% β-mercaptoethanol (βME; Bio-Rad, Hercules, CA, USA) with acetone (Fisher Scientific, Pittsburgh, PA, USA; volume:volume). A 10.0% trichloracetic acid (TCA; Sigma, St. Louis, MO, USA)/acetone 'suspension solution' was prepared using 300mL of the stock solution. The remainder of the stock solution was used as a 'wash solution' to remove TCA and impurities from the sample after suspension. The solutions were chilled in a -20˚C freezer for a minimum of one hour.

Approximately 2.0g of termites (400 termites ≈ 1.0 gram) were harvested from a laboratory stabilized termite colony. The termites were transferred into a mortar pre-chilled to -20˚C, covered with liquid nitrogen, and ground to a fine powder. Sub-freezing conditions of the termite sample were maintained by adding liquid nitrogen to the mortar as needed. A pre-chilled (-20˚C) 50mL Beckman (Beckman Coulter, Inc., Fullerton, CA USA) centrifuge tube was filled with 10mL of suspension fluid and the resulting termite powder transferred from the mortar into the 50mL centrifuge tube using a metal spatula. The spatula and tube sides were rinsed into the tube with an additional 20mL of suspension solution. The termite protein sample was then incubated for at least one hour at -20˚C.

An empty Beckman centrifuge tube was filled with water to the same total weight of the centrifuge tube containing the protein sample. This provided a counter-balance for the protein sample in the centrifuge. A JA-17 rotor (Beckman Coulter, Inc., Fullerton, CA, USA) was loaded into a Beckman J2-HS Centrifuge and the counter-balanced

centrifuge tubes were placed into opposing tube receptacles in the rotor. The samples were centrifuged at 16,000 rpm at 0˚C for 30 minutes. Both centrifuge tubes were removed from the rotor and returned to the lab. The supernate was removed from the centrifuge tube containing the protein pellet and discarded. A wash cycle was performed by using wash solution to wash the pellet for 2.0 minutes. Next, the wash solution was decanted and discarded. The remaining protein sample was then covered with 5.0mL of wash solution. The counter-balance tube was rebalanced to the same weight as the sample centrifuge tube. The counter-balanced centrifuge tubes were returned to opposing tube receptacles in the rotor and centrifuged at 16,000 rpm at 0˚C for 15.0 minutes to complete the wash cycle. The wash cycle was repeated a minimum of three times until the removed supernate was clear. When the removed supernate was clear, the protein sample was washed a final time and centrifuged at 16,000 rpm at 0˚C for 1.0 hour. The supernate was discarded and the protein extract pellet transferred to multiple microcentrifuge tubes to expedite drying the protein sample. The protein extract was dried under nitrogen and stored at -80˚C.

Whole-body protein extraction with protease inhibitor – The protocol for a whole-body protein extraction was repeated, but added 300µl 'protease inhibitor cocktail for general use' (Sigma P2714) for each 30mL of wash solution prepared.

Protein solubilization using urea – A 200mL reswelling solution of 8.0M urea and 2.0% Triton X-100 (GE Healthcare, Pittsburgh, PA, USA) in ddH_2O was prepared. The solution was divided into 1.6mL aliquots and stored at -20˚C.

The manufacturer's recommended volume (Table 1) of 60mM dithiothreitol (DTT) IPG reswelling solution for the IPG strip being reswelled was added to a 2.0mL

microcentrifuge tube. For each 300µl of resolubilized protein solution prepared, 10.0mg of protein extract was added to the IPG reswelling solution. The resulting protein solution was chilled for 30 minutes in a bucket of ice. The chilled tube was transferred to a beaker of ice and ultrasonicated using a Fisher Sonic Dismembrator Model 300. The Sonic Dismembrator had the small tip installed and was operated at 35% power for 20 seconds. The sample was chilled in ice for 4.0 minutes after sonification. This ultrasonification/chilling cycle was repeated three times, for a total period of 1 minute and 20 seconds. After ultrasonification was completed, a 2.0% volume of ampholyte with a pH range matching the IPG strip pH range was added to the microcentrifuge tube. The tube was closed and placed on the rocker platform having an oscillation rate of 1–3 and incubated at room temperature for 1.0–2.0 hours. The tubes were removed from the rocker and counter-balanced before being placed into opposing rotor wells of a ProFuge^{tm} 10K centrifuge (Stratagene, La Jolla, CA, USA). The tubes were centrifuged for 60 seconds at 10,000rpm to clarify the protein solution. Protein concentration was checked using a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and/or Ramagli's Protein Assay (Ramagli 1999). The ND-1000 is a full-spectrum UV-Vis absorbance analyzer capable of analyzing sample volumes as low as 1µl. It can be used for measuring absorbance in DNA, RNA, and proteins (NanoDrop Technologies 2006). Protein concentration was determined using the A280 assay analysis module of the ND-1000 software. This module does not require a standard curve and uses absorbance measured at 280nm to calculate protein concentration (mg/mL).

Protein solubilization using urea and CHAPS – The 'Protein Solubilization' protocol was repeated, but adding a 2.0% amount (w:v) of 3-[(3-Cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS) in ddH_2O (Boonmee et al. 2003) to the initial protein solubilization solution.

Protein solubilization using urea, CHAPS, and tris – The 'Protein Solubilization' protocol was repeated, but substituting the stock reswelling solution of 8.0M urea and 2.0% Triton X-100 in ddH_2O with a stock reswelling solution of 8.0M urea, 4.0% CHAPS, and 40mM Tris in ddH_2O (Hernandez et al. 2004).

Protein solubilization using urea and thiourea – The 'Protein Solubilization' protocol was repeated, but substituting a reswelling solution of 6.0M urea, 2.0M thiourea, and 2.0% volume Triton X-100 in lieu of the reswelling solution of 8.0M urea, 2.0% thiourea, and 2.0% volume of Triton X-100 in ddH_2O.

Protein solubilization using urea, thiourea, and CHAPS – The 'Protein Solubilization' protocol was repeated, but substituting the stock reswelling solution of 8.0M urea and 2.0% Triton X-100 in ddH_2O with a stock reswelling solution of 6.0M urea, 2.0M thiourea, 2.0% CHAPS, 60mM DTT, and 2.0% Triton X-100 with in ddH_2O (Rodriguez-Ortega et al. 2003).

Protein solubilization using urea, thiourea, and tris – The 'Protein Solubilization' protocol was repeated, but substituting the stock reswelling solution of 8.0M urea and 2.0% Triton X-100 in ddH_2O with a stock reswelling solution of 6.0M urea, 2.0M thiourea, 2.0% Triton X-100, and 20mM Tris with in ddH_2O (Sprenger et al. 2004).

Protein solubilization using urea, thiourea, CHAPS, and tris – The 'Protein Solubilization' protocol was repeated, but substituting the stock reswelling solution of

8.0M urea and 2.0% Triton X-100 in ddH_2O with a stock reswelling solution of 6.0M urea, 2.0M thiourea, 2.0% CHAPS, 2.0% Triton X-100, and 40mM Tris with in ddH_2O (Musante et al. 1998).

Ramagli's protein assay – A 10.0mL stock reswelling solution (8.0M urea and 2.0% Triton X-100 in ddH_2O) containing 60mM dithiothreitol (DTT) and 2.0% volume of ampholyte was prepared. The stock solution was divided into two 5.0mL aliquots. A 'standard' solution of ovalbumin with a final protein concentration of 5.0µg/µl was prepared. Increasing volumes (0, 1, 2, 3, 4, 5, 6, 8, and 10µl) of ovalbumin solution were added to a series of 0.6mL microcentrifuge tubes. Adequate reswelling solution was added to bring each tube to 10.0µl total volume. Next, 10.0µl hydrochloric acid (HCl) and 80.0µl ddH_2O were added to each tube. These known samples were used to generate a standard curve for evaluating the protein quantities in the unknown samples. This standard protocol was repeated using unknown protein samples. To determine the best dilution for the unknown protein(s), 1.0, 5.0, and 10.0µl volumes of the protein solution having an unknown concentration were used.

Protein assay dye solution was prepared by adding 4.0mL of Bio-Rad Protein Assay Dye Reagent Concentrate to 16.0mL ddH_2O in a plastic container. Using a plastic 96-well plate, 20.0µl from each standard and unknown protein sample was added to individual wells. Adequate protein assay dye solution was added to each well containing a protein sample to bring the volume to 200.0µl (Ramagli 1999) and incubated for 5.0 minutes. The 96-well plate was read on a Bio-Rad Model 3550 Plate Reader using a wavelength of 595nm. Microplate Manager® version 4.01beta was used to record the

data, generate the standard curve, and analyze the unknown protein concentrations. The standard curve graph, the data table, and the unknown concentration report were printed.

Coomassie brilliant blue staining – A 0.025% (w:v) Coomassie brilliant blue (CBB) dye was prepared by adding CBB R-250 powder to a solution comprised of 40.0% methanol, 7.0% acetic acid, and 53.0% ddH_2O and filtered using a Corning® 0.45µl cellulose acetate filter. A CBB destain solution was prepared by combining 50.0% methanol, 38.0% ddH_2O, and 12.0% acetic acid (Hoefer Scientific Instruments 1994).

Gels were removed from the plates and transferred into staining trays. Gels were rinsed by covering the gels with ddH_2O, then immediately decanted. Next, the gels were covered with 0.025% CBB dye, and were allowed to absorb dye for 1.0 hour. The dye was removed and the gels were rinsed by covering the gels with ddH_2O, then the water was immediately decanted. The gels were then covered CBB destain solution, adding two (2.0 × 10.0 × 1.5cm) pieces of closed-cell foam to the tray. The gels were allowed to destain for 2.0 hours, replacing the foam if it became saturated with CBB stain. Gels were removed from CBB destain solution and rinsed in ddH_2O prior to storage in 5.0% acetic acid solution. Gel storage was maintained in 5.0% acetic acid solution at 4°C (Berkelman and Stenstedt 1998).

Silver staining – Gels were removed from the plates and transferred into staining trays, then rinsed with ddH_2O and placed into fixing solution (50.0% methanol, 12.0% acetic acid) for a minimum of 90 minutes. The fixing solution was decanted and gels were washed twice in 50.0% ethanol combined with 50.0% ddH_2O for 20 minutes. Next, gels were washed in ddH_2O for 20 minutes. Gels were then treated with a 0.2g/L sodium thiosulfate ($Na_2S_2O_3•5H_2O$) solution for 1.0 minute and rinsed three times with ddH_2O

for 20 seconds each time. Next, gels were impregnated with 2.0g/L silver nitrate ($AgNO_3$) for 30 minutes at 4.0˚C, followed by three consecutive 20-second rinses with ddH_2O. Gels were transferred to a clean staining tray and rinsed with ddH_2O for 1.0 minute. Then, gels were developed for approximately 10 minutes in silver stain developer (60.0g/L sodium carbonate, 4.0mg/L sodium thiosulfate, 0.5m/L 37% formalin), removed from silver stain developer, and rinsed for 5 seconds in ddH_2O. Gels were then washed in 5.0% acetic acid for 2.0–3.0 minutes to stop development, removed and rinsed in ddH_2O. Gels were stored in 5.0% acetic acid until spots were harvested for mass spectrometry analysis (Blum et al. 1987).

Part 1 overview – Each potential solubilization protocol was evaluated using Amersham Biosciences (now part of GE Healthcare) 4–7pH IPG strips and 10–20% acrylamide gradient pre-cast Ready Gels (7.0 × 8.3cm; 1.0mm thickness). Visualization was completed with coomassie brilliant blue as well as silver stain. The protocol that yielded the highest resolution and the most complete protein map was selected as the optimized solubilization protocol.

The optimized solubilization protocol was repeated using an Amersham 4–7pH IPG strip and a 10–20% Ready Gel. However, visualization was completed with silver nitrate staining. The resulting gel was compared with the gel selected from the optimized solubilization protocol visualized using Coomassie brilliant blue. The gel that yielded the highest resolution gel with the most complete protein map was selected as the optimized staining protocol.

Part 2: Optimize isoelectric focusing

Reswelling IPG strips – A trace (2–4 crystals) of bromophenol blue was added to the protein solution prepared in 'Part 1: Optimize solubilization', then stained and mixed. The Amersham reswelling tray was placed on a flat surface and the legs adjusted until the bubble-level was centered. The protective cover from the gel side of the IPG strip was removed and the IPG strip rinsed with ddH_2O. Excess water was removed from the strip by drawing its back side across a clean piece of filter paper. The proper amount of dilute protein solution, for the pH range and staining protocol (Table 1) tested, was placed into a reswelling tray well, with the gel side in contact with the protein solution that was evenly distributed beneath the strip. The strip was then covered with 2.5mL of glycerol to prevent evaporation and urea crystallization, and reswelled for a minimum of 10 hours.

Table 1. Immobiline pH strip protein loads and DDT reswelling solution volume

IPG Strip		Protein Load		Minimum
Length (cm)	pH range	Silver (µg)	Coomassie blue (µg)	Volume Per Strip (µl)
7	3–10, 3–10NL	2–4	200–400	125
7	4–7	4–8	400–800	125
11	3–10, 3–10NL	4–8	400–800	200
11	4–7	10–20	1,000–2,000	200
24	3–10, 3–10NL	20–45	2,000–4,500	450
24	4–7	45–90	4,500–9,000	450

Isoelectric focusing – A Fisher Scientific Isotemp Refrigerated Circulator Model 900 (chiller) was used to circulate refrigerated water through a Multiphor™ II Electrophoresis System to maintain the cooling plate at 10.0°C. The Multiphor™ Electrophoresis System tray was connected to the chiller and the cooling plate leveled. Using a 10mL glass pipette, distribute approximately 12mL of DryStrip cover fluid (mineral oil) in an oval shape onto the cooling plate. The electrode tray was placed onto the cooling plate so the

mineral oil was evenly dispersed without visible air bubbles between the electrode tray and cooling plate. A 12mL volume of mineral oil was placed into the electrode tray and the alignment tray was placed into bottom of the electrode tray so the mineral oil was dispersed evenly without visible air bubbles in the oil. Air bubbles remaining between the alignment tray and the electrode tray were gently pressed out.

Two pieces of IPG Electrode Strip material (110 × 6mm) were prepared by placing the strips on a piece of glass and then evenly distributing 0.5mL of ddH_2O along the length of each strip and put aside wait until needed. Next, an IPG strip was removed from the reswelling tray well and rinsed thoroughly with ddH_2O. Excess water was removed from the strip by drawing its back side across a clean piece of filter paper. The IPG strip was placed gel side up into an alignment tray groove with the anodic end of the strip toward the anode (red) end of the electrode tray. This procedure was repeated until all previously prepared IPG strips were placed into the alignment tray.

Excess moisture was removed from the two pieces of previously prepared electrode strip material by patting each once with a piece of filter paper. Electrode strip material was then placed perpendicularly across each end of the IPG strip(s) so approximately 3mm of the electrode strip width rested on the IPG strip gel surface. Electrodes were placed along the center of the electrode strip material: the red electrode to the anode end and the black electrode to the cathode end of the IPG strip, respectively.

The alignment tray and strips were covered with 80–90mL of Immobiline™ DryStrip cover fluid. Air bubbles forming in the cover fluid were removed and the Multiphor lid secured before the system was connected to the power supply. The power supply was programmed to operate under constant voltage conditions in two stages:

1) 6 hours @ 200V and, 2) 18 hours @ 3500V. Upon completion, the IPG strips were removed from the system and rinsed with ddH_2O. Excess water was removed and the IPG strips were stored at -80˚C or used immediately (Berkelman and Stenstedt 1998).

Part 2 overview – Isoelectric focusing (IEF) was optimized by evaluating IPG strips with pH ranges of 3–10, 3–10 non-linear (NL), and 4–7, using the optimized solubilization protocol. The second dimension was performed using 10–20% Ready Gels. Visualization was performed using the optimized staining protocol. The gel yielding the highest resolution with the most complete protein map was selected as the optimized IEF protocol.

Part 3: Optimize SDS-PAGE

Gel fabrication – A 30% acrylamide monomer was prepared by dissolving acrylamide into ddH_2O. One of two crosslinkers was added depending on the requirements: 0.8% N,N^1-methylene bis-acrylamide (Amersham Biosciences, Uppsala, Sweden) or 0.8% piperazine diacrylamide (Bio-Rad, Hercules, CA). The resulting solution was filtered using a Corning 0.45µm cellulose acetate low protein binding membrane filter system and stored in a dark bottle at 4.0°C in zero light conditions.

Glass plates (276 × 216.5 × 0.825mm) were thoroughly cleaned before assembly in a casting apparatus. Two glass plates were separated by 1.5mm spacers placed along each vertical edge to form gel cassette. A modified Laemmli protocol (Laemmli 1970) was used to prepare a separating gel solution at the desired concentration by combining water, tris-HCL (pH 8.8), and 30% acrylamide monomer solution in a flask. This solution was degassed for 15 minutes by sonicating the solution while under a vacuum. Appropriate volumes (Table 2) of 10.0% SDS, 10.0% ammonium persulfate, and

TEMED were then added. Immediately after combining the separating gel reagents, the solution was poured into the gel cassette void until the solution surface was within 2mm of the top of the glass. Water-saturated butanol was gently applied atop the separating gel surface to facilitate a flat IPG strip interface. The casting apparatus was covered with plastic wrap to minimize evaporation. After 1.0 hour, the butanol was decanted and replaced with ddH_2O. Gels were allowed to polymerize overnight. Gels were either used immediately or placed into storage solution for short-term storage at 4.0°C.

Table 2. Reagent volumes for acrylamide solutions per each 10.0mL of solution

Reagent	12.0%	14.0%	16.0%	18.0%
Distilled H_2O	3.35mL	2.68mL	2.02mL	1.35mL
1.5M Tris-HCl (pH 8.8)	2.50mL	2.50mL	2.50mL	2.50mL
Acrylamide/Bis (30% stock)	4.01mL	4.67mL	5.33mL	6.00mL
10.0% (w/v) SDS stock	100.00µl	100.00µl	10.000µl	100.00µl
Fresh 10.0% ammonium persulfate	50.00µl	50.00µl	50.00µl	50.00µl
TEMED	5.00µl	5.00µl	5.00µl	5.00µl

Gel electrophoresis – Water was added to a VWR Scientific Heat Block (VWR International, West Chester, PA, USA) and heated to 60˚C. Melted 1.8mL agarose aliquots were placed in the heat block and allowed to cool to 60˚C. Next, previously prepared IPG strips were rinsed with ddH_2O, excess water removed, and placed in an incubation tube. SDS equilibration buffer with 1.0% DTT was added to the incubation tube and the IPG strips were allowed to equilibrate for 15 minutes. Following incubation, the strips were rinsed and the IPG strips were equilibrated for an additional 15 minutes using 2.5% (weight:volume) iodoacetamide in SDS equilibration buffer.

While the IPG strips were equilibrating, 1.0× and 2.5× running buffer solutions were prepared by diluting a 10.0× SDS running buffer. The1.0× running buffer solution

was added to the lower buffer chamber of the electrophoresis and circulated to facilitate air bubble removal. The electrophoresis unit was connected to a Caron (Marietta, OH, USA) Refrigerated and Heated Circulating Bath (chiller) and the running buffer was chilled to approximately 12°C.

After equilibration was completed, the gel well was rinsed with ddH_2O. Excess water was removed from the gel well by inserting filter paper into the void above the gel between the plates. The IPG strip was seated into the gel well so the cathodic (-) end was against the left plate spacer and no air bubbles remained between the strip and gel surface. Three electrode wicks (Bio-Rad Catalog # 1654071) were placed along the edge of the gel and a total of 20.0µl of Amersham Full Range Rainbow Molecular Weight Markers (2007 Catalog # RPN800) was applied to the combined wicks. The gel well was filled with 60.0°C agarose using a disposable glass pipette. The gel cassettes were placed into the lower tank of the electrophoresis unit and empty cassette slots were filled with acrylic cassette blanks. Adequate 1× running buffer was added to the lower tank to bring the level up to the fill line.

The upper buffer tank was placed over the lower buffer tank and filled with 2.5× running buffer. The tank was covered and the power leads connected to an Amersham Biosciences EPS 601 power supply. The power supply was programmed for two stages: Stage 1 was set for 1.0 hour at 15mA per gel; Stage 2 was set for 24.0 hours at 40mA per gel. The second stage was stopped when dye front reached approximately 2.0 mm from the gel bottom. The gel cassettes were removed from tank and gels transferred to staining trays. Gels were stained using either silver stain or Coomassie brilliant blue staining techniques (Berkelman and Stenstedt 1998).

Part 3 overview – The SDS-PAGE protocol was optimized by evaluating a series of hand-cast gels (7.0 × 8.3cm; 1.5mm thick) containing ascending acrylamide concentrations of 12.0, 14.0, 16.0, and 18.0%. Criterion gradient gels (13.3 × 8.7cm; 1.0mm thickness) were also evaluated for the 8–16% and 10–20% acrylamide ranges. The protocol yielding the highest resolution with the most complete protein map was selected for use in the optimized SDS-PAGE protocol. The optimized solubilization, IEF, SDS-PAGE, and staining protocols were employed in subsequent objectives.

Objective II – Develop standard protein reference map for *Reticulitermes flavipes*

Image acquisition – Stained gels were converted to digital format with an Epson Expression 1680 Scanner interfaced with Epson Scan (version 1.01E) for Microsoft Windows (Redmond, WA, USA). Each gel was scanned at 300dpi as both a 24-bit color and 12-bit grayscale image. Each image was saved as an uncompressed TIFF file and named using the following criteria: Termite_Colony <#>_Rep <#>_<IPG strip #>_<pH range>pH_<gel acryalamide %>_<solubilization protocol>_<IPG length>_<series image #>.tif. For example, a filename might be:

Termite_Colony 1_Rep12_#54321_3–10pH_16%_6MU2.0MT_24cm.tif

Image analysis – Image analysis was undertaken using Synoptics Dymension 2 gel analysis software version 2.0.6.7. Uncompressed TIFF gel images were opened into Dymension 2 for analysis. The Dymension software was configured to optimize the spot analyses of the gels. "Configure Background Correction For Experiment" was set to 'radius fraction' equals three and 'intensity fraction' equals 20. Each gel was normalized to remove background noise allowing protein quantification to be attempted for each spot. The "Patch Analyzer" tab of the "Spot Detector Settings" was adjusted to: 'blur radius' equals 1.4, 'detection confidence' equals 40, 'separation confidence ratio' equals 27, 'peak limit threshold' equals 0.005, 'splitting threshold' equals 0.008, and 'split spots automatically was checked. Each gel was analyzed for spot identification. Manual splitting and merging of spots was completed where automatic detection was not adequately performed by the software. Ladders correlating to pI and molecular weight were added allowing each spot to be identified on a Cartesian plane with the x-dimension corresponding to pI and the y-dimension corresponding to molecular weight. Secondary

gels were warped by distending them until corresponding spots were located in approximately the same location. This facilitated visual confirmation of the spot matching between and among gels. Spot matches between and among gels were identified and recorded for comparative protein quantification analysis. A protein gel map was prepared by selecting the gel having the most ideal regions displaying the optimal distribution of spots. The gel background was removed using the automatic background removal function in Dymension 2, then the image was exported as a TIF file. Photoshop CS2 (Adobe Systems Incorporated, San Jose, CA, USA) was used to adjust contrast and brightness to facilitate printing. Spots processed for mass spectrometry were circled on the gel map using Microsoft Powerpoint. Each circled spot was assigned a numeric identifier. Mapedit (Boutell.com, Inc., Philadelphia, PA, USA) was used to create a digital representation of the map for access via the "2D Gel Analysis" webpage, Oklahoma State University Entomology and Plant Pathology website. This page is located on the faculty page of Jack Dillwith, Professor (OSU Entomology and Plant Pathology 2006). Mapedit facilitated HTML mapping of individual spots allowing initial spot data to be accessed by mouse rollover and additional data to be accessed by mouse click.

Objective III – Initial characterization of the putative *R. flavipes* proteome

Protein digestion – Microcentrifuge tubes (600µl) were washed using freshly prepared solution of 0.1% triflouroacetic acid (TFA) in 50% acetonitrile (ACN). Samples were harvested by excising 1 to 3 plugs from a single spot with a 1.5mm diameter plastic straw rinsed with 0.1% TFA in 50% ACN. Samples were placed into a washed 600µl tube and rinsed with ddH_2O. A wash solution containing 25mL ammonium bicarbonate (ABC) and 25mL ACN was prepared. Wash solution was added to tubes until samples are covered. Tubes were placed on rocker and incubated for 2 hours. Wash solution was replaced and rocked for an additional two hours, and repeated for a total of three washes. Wash solution was removed from samples. Samples were then rinsed with 100% ACN and dried by centrifugation in a 0 Torr vacuum for 30 minutes in a Savant Speed-Vac Concentrator with a Refrigerated Condensation Trap (Global Medical Instrumentation, Inc., Ramsey, MN, USA; speed vac).

Adequate 0.25µg/µl trypsin solution was added to each tube until samples were submerged. Tubes were incubated at room temperature for 30 minutes. If gel samples did not remain covered, 5µl aliquots of trypsin solution were added until all samples were covered. Tubes were incubated at 37˚C overnight (16 hour minimum) in a MJ Research, Inc. (Bio-Rad) PTC-100 Programmable Thermal Controller (Shevchenko et al. 1996).

Peptide extraction – After tubes were removed from the thermal controller, 20µl of 50% ACN/0.1% TFA were added to each tube. The solution was allowed to extract peptides for 1.5 hours. Next, the peptide solution was transferred to a ‘collection’ tube previously rinsed with 0.1% TFA in 50% ACN. The peptide extraction was repeated two times using 10µl of 0.1% TFA in 50% ACN. The peptide solution was concentrated by

evaporating the ACN solution in a non-heated Speed Vac to a final volume of 10–15µl (Jensen et al. 1999).

Protein identification –Matrix solution was prepared by saturating 0.1% TFA/50% ACN with recrystallized alpha-cyano-4-hydroxycinnamic acid (CHCA) (Hartson et al. 2003). Protein identification was attempted by identifying digested protein sample using mass spectrometry. Mass spectrometry (MS) was conducted using an Applied Biosystem Pro-Star MALDI-TOF instrument controlled by Voyager version 5 with Data Explorer software (Applied Biosystems, Foster, CA, USA).

Saturated CHCA solution was centrifuged for 30 seconds to remove any non-solubilized material out of solution. Protein samples were spotted onto a 96×2 well stainless steel MALDI plate using a 0.8µl sample size. Equivalent volumes of CHCA solution were layered over the protein sample and allowed to dry (Bienvenut et al. 2002). The samples were analyzed using MALDI-TOF with the following instrument settings: reflector mode, positive ion mode, 20,000V acceleration voltage, 75V grid voltage, 0.02V guide wire voltage, and 300nsec delay time. Data from 200 TOF analyses were accumulated for each sample (Hartson et al. 2003).

Bioinformatics – Oklahoma State University Bioinformatics Facility maintains numerous databases locally for university use. These databases are updated at least monthly and include databases from NCBInr, MSDB, UniRef100 and SPROT. Additionally, the Bioinformatics facility annotated and maintained an expressed sequence tag (EST) database named “EST_others”. This database contained sequences from over three million arthropod peptides and was updated weekly.

Proteins were identified using Mascot® Daemon version 2.1.0 (Matrix Science, Boston, MA, USA) to search proteins within various databases. The databases used were SwissProt 50.4 (5,858 arthropod entries), NCBInr 20060825 (957,081 metazoa entries), Uniref100 20060725 (914,348 metazoa entries), and EST_others 20051027 (683,680 arthropoda entries). Mascot Daemon used Data Explorer functions to deisotope spectra and generate a peak list from 500*m/z* to 3000*m/z* using the five most intense peaks from each 100*m/z* interval. The peak list was searched against the selected database. Other search parameters included the following settings: no fixed modifications; variable modifications of carboxymethyl (c), oxidation (M), propionamide (c), and pryo-glu- (N-term Q). Additional settings were monoisotopic peaks; protein mass equals kDa, Peptide tolerance equals 100ppm, and peptide charge equals '1+'. Data import filter was set to "Data Explorer (Voyager) MS Data". A contaminant list was used to automatically filter experimental peaks generated by known masses such as trypsin, keratin, and acrylamide. Tasks were submitted with the following 'task name' format: <last name> <IEF serial number> <database name> <taxonomy> <date>. An example of a task name would be: Bowen 05743 NCBInr Arthropoda 22 AUG 2006.

Molecular function – Protein functions were estimated by matching termite proteins to protein database entries having known functions. Mass spectrometry data base results were sent to the Oklahoma State University Bioinformatics group in the Department of Biochemistry and Molecular Biology. Patricia Ayoubi conveyed the following protocol used for conversion of protein data to gene data to protein functions. "The accession numbers of the most significant matches by Mascot were used to obtain the corresponding amino acid sequences in FASTA format. Each peptide sequence was then

compared to the UniProt database (http://www.pir.uniprot.org/) using the BLASTP amino acid alignment program (Altschul et al. 1990) and the best match with an e-value of 1e-20 or less was obtained. Finally, the predicted molecular function of the analyzed proteins was then inferred from Gene Ontology term assignments (Camaschella et al. 2000) to these matching UniProt records. Assignment of GO terms to each UniProt record was obtained from annotations of the multispecies European Bioinformatics Institute electronic annotation and assignment to Gene Ontology terms to the UniProt Knowledgebase (release 111092006 posted at http://cvsweb.geneontology.org/cgi-bin/cvsweb.cgi/go/gene-associations/gene_association.goa_uniprot.gz?rev=HEAD). After GO term assignments were established for the analyzed proteins from termites, the distribution and proportion GO terms molecular functions was then determined."

Chemically assisted fragmentation – Protein identification was also attempted on selected protein spots using post source decay (PSD) of the protein peptides using protocols from Marekov and Steinert (2003) and Chen et al. (2004) to confirm protein identifications. Chemically assisted fragmentation (CAF) was undertaken at the Biochemistry and Molecular Biology (BMB) Core Facility located at Oklahoma State University. Sequence tags were generated using sulfonated peptides and these were used to manually calculate peptide sequences. The peptide sequences were searched against the NCBInr and SwissProt databases using the NCBI BLAST (National Center for Biotechnology Information 2006) "Search for short, nearly exact matches" option under the 'Protein' tool.

Tandem mass spectrometry – Protein samples were sent to the Nevada Proteomics Center located on the University of Nevada, Reno. Rebakah Woolsey was the technician and conveyed the following protocols used for protein digestion and mass spectrometry:

Protein digestion – "Spots are digested using Investigator^{TM} Proprep^{TM} (Genomic Solutions, Ann Arbor; MI), using a previously described protocol (Rosenfeld et al. 1992) with some modifications. Samples are washed twice with 25mM ABC and 100% acetonitrile, reduced and alkylated using 10mM DTT and 100mM Iodoacetamide and incubated with 75ng Trypsin in 25mM ABC for 6.0 hrs at 37^{o} C.

Samples are prepared and spotted onto a MALDI target with ZipTip$^{®}$ μ-C18 (Millipore, Billerica; MA). Samples are aspirated and dispensed 3 times and eluted with 70% ACN, 0.2% formic acid and overlaid with 0.5μl 5mg/mL MALDI matrix (α-Cyano-4-hydroxycinnamic acid) and 10mM ammonium phosphate."

Tandem mass spectromtry – "All mass spectrometric data were collected using an ABI 4700 MALDI TOF/TOF (Applied Biosystems, Foster City; CA). Data were acquired in reflector mode from a mass range of 700–4,000 Daltons (Da) and 1,250 laser shots were averaged for each mass spectrum. Each sample was internally calibrated on trypsin's autolysis peaks. The eight most intense ions from the MS analysis, which were not on the exclusion list, were subjected to MS/MS. For MS/MS analysis the mass range was 70 to precursor ion, with a precursor window of -1 to 3 Daltons with an average 5,000 laser shots for each spectrum. Data were stored in an Oracle database (Oracle Corporation, Redwood Shores, CA, USA).

Data were extracted from the Oracle database and a peak list was created using IDQuest software (Bio-Rad Laboratories) from the raw data generated from the ABI 4700

MALDI TOF/TOF. This peak list was based on signal to noise filtering and an exclusion list and included de-isotoping. The resulting file was then searched by Mascot™ (Matrix Science, Boston; MA). A tolerance of 20ppm was used if the sample was internally calibrated, or 200 ppm tolerance if the default calibration was applied. Database search parameters include 1 missed cleavage, oxidation of methionines, and carbamidomethylation of cysteines."

Objective IV – Test for differential protein expression among *R. flavipes* colonies

Differential comparison – Using the optimized protocols from Objective 1, optimized gels were generated from the second and third colonies. Additionally, gel replicates for each colony were generated using the optimized protocols. Gel replicates were typically generated in groups consisting of one gel for each colony and were accrued throughout the year. Gels were scanned and analyzed as defined in Objective 2. Using Dymension, gel images of three replicates from each of the three colonies were selected for comparison. The selected gel images were loaded into Dymension software as an 'experiment' and the gel images for each colony were treated as a 'sample', yielding two samples named sample 1– 3. These samples were renamed Colony 1, Colony 2, and Colony 3, respectively. The first gel image in each sample was used as the warping reference and each subsequent gel image was warped (stretched) to facilitate matching spot patterns to the reference image. Then each matching spot was indicated on the sample warping reference as a consensus spot. Once each sample was compared, the samples were compared against the first sample. When sample matching was undertaken, the first gel image of the first sample became the warping reference for all the gel images in all the samples. Each matching spot among the samples are indicated as consensus spots, while non-matching spots are indicated as non-consensus spots. Additionally, manual comparison of gel images was undertaken. Images were divided into roughly equal thirds and equivalent thirds were aligned next to each other to simplify visual comparison.

Spot patterns from each colony were compared for differences in protein expression. Differential proteins were selected and extracted for identification. Mass

spectrometry and protein identification were completed using the same methods defined in Objective 3.

Objective V – Test for differential protein expression between worker and solder castes

Termite collection – Soldiers were harvested from Colony 1. Because 1.0g of soldiers was collected, the protocols were scaled down proportionally.

Differential comparison – Using the optimized protocols from Objective 1, optimized gels were generated from the soldier termites. The gels were scanned and analyzed as defined in Objective 2. Additionally, gel images of three replicates from each of the worker and soldier castes of Colony 1 were selected for comparison. The selected gel images were loaded into Dymension software as an 'experiment' and the gel images for each caste were treated as a 'sample', yielding two samples named Sample 1 and Sample 2. These samples were renamed Worker and Soldier, respectively. The first gel image in each sample was used as the warping reference and each subsequent gel image was warped (stretched) to facilitate matching spot patterns to the reference image. Then each matching spot was indicated on the sample warping reference as a consensus spot. Once each sample was compared, the samples were compared against the first sample. When sample matching was undertaken, the first gel image of the first sample became the warping reference for all the gel images in all the samples. Each matching spot among the samples are indicated as consensus spots, while non-matching spots are indicated as non-consensus spots. Using Dymension, data from the soldier caste gels were compared against the Colony 1 data (worker caste) for similarities in protein expression.

Voucher specimens

Voucher specimens were collected from each colony and reposited in OSU's K.C. Emerson Entomology Museum.

RESULTS AND DISCUSSION

Objective I – Develop a system for analyzing the termite proteome

Screening system. A basic system was developed using components available in the laboratory and used to optimize components of the screening system. Protein extraction and solubilization were undertaken using methods adapted from earlier non-related projects. The first dimension was completed using Amersham 4–7pH Immobiline pH Gradient Drystrips. The second dimension was completed using the Bio-Rad Ready Gel system. This system provided capability to cast custom gels, as well as use a wide variety of commercially-purchased gels having acrylamide gradients. Initial screening was undertaken using 10–20% acrylamide gradient gels. Spot visualization was accomplished using both Coomassie brilliant blue dye and silver nitrate staining.

Part 1: Optimize solubilization and staining

System development: Extraction of proteins from termites. Protein extraction began during the grinding of termite tissue. TCA-acetone extraction was selected as the method to facilitate precipitation of proteins from the termite powder. This method is recognized as an effective means of removing impurities and interfering substances during tissue preparation (Simpson 2003).

System development: Determine usage of protease inhibitor. Whole-body termite precipitant was prepared using protease inhibitor, and without protease inhibitor. When cells are disrupted and release proteases, proteins can be exposed to random proteolytic

activity. This activity could potentially make the protein spot pattern unpredictable. To prevent proteolysis, reagents such as 8.0M urea or 10.0% TCA are used to minimize proteolytic effect. Other methods to reduce or prevent proteolysis would be

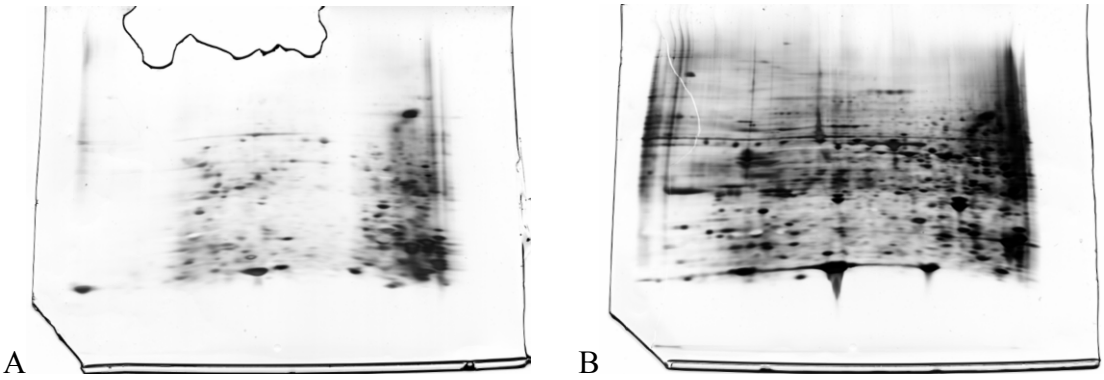


Fig. 1. Examples of proteins processed with and without protease inhibitor: (A) Precipitant processed with protease inhibitor; (B) Precipitant processed without protease inhibitor.

preparation of samples in a low temperature environment, slowing proteolysis or the addition of a protease inhibitor to the protein solution (Berkelman and Stenstedt 1998). The proteins prepared using protease inhibitor exhibited a better defined spot pattern (Fig. 1A). The gel containing proteins prepared without protease inhibitor exhibited more protein spots (Fig. 1B), but the results were less consistent and more difficult to reproduce. The addition of protease inhibitor minimized random cleavage of proteins and proteolytic activity in the solution containing whole body precipitant. The use of protease inhibitor was selected as the best choice for this stage of the optimized system.

System development: Optimization of solubilization solution. Seven methods of solubilizing proteins from the TCA-acetone precipitant were identified from the literature. Three of the methods were based on an 8.0M urea solution (Fig. 2), while four were based on a solution of 6.0M urea and 2.0M thiourea (Fig. 3). Urea is a chaotropic agent used to help solubilize proteins by disrupting hydrogen bonding. The addition of thiourea increases the chaotropic function of the solution which should enhance the

solubility of some proteins. Triton X-100 is a non-ionic surfactant and acts to disrupt hydrophobic interactions. Since it is non-ionic, it does not interfere with the charge on the proteins during IEF. This reagent can also facilitate protein solubility at the protein pI. Carrier ampholyte is necessary for isoelectric focusing. It is a solution containing both acid and base buffers and should be matched to the same pH range as the

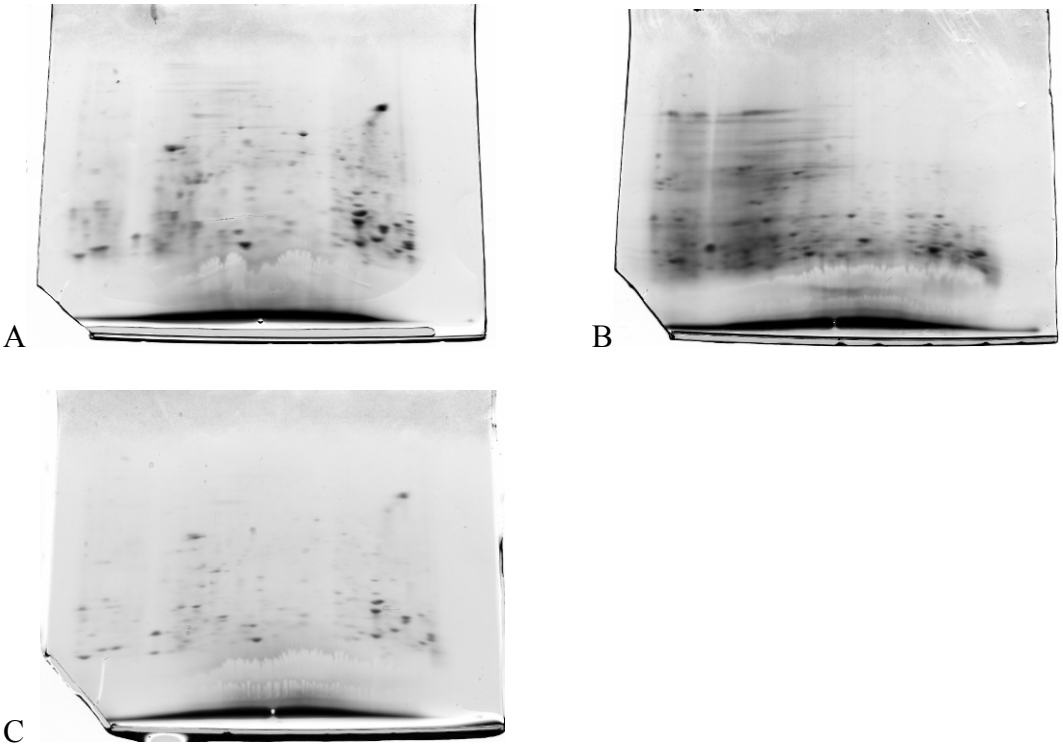


Fig. 2. Examples of gels using 8.0M treatments: (A) 8.0M urea, Triton X-100, and carrier ampholyte; (B) 8.0M urea, Triton X-100, tris, and carrier ampholyte; (C) 8.0M urea, Triton X-100, CHAPS, and carrier ampholyte.

immobiline pH gradient strip. Ampholyte acts as a salt in the solubilized protein solution preventing possible protein precipitation as salts are removed by the high current first stage of isoelectric focusing (Bio-Rad 2002).

The seven solubilization methods were conducted as described in 'Materials and Methods' (pp. 20–23) and used to prepare gels for comparison. Three gels using 8.0M urea solutions were prepared first and the best method identified. Of these gels, the 8.0M

urea solution with only Triton X-100 and ampholyte added (Fig. 2A) produced a spot pattern with high resolution and low streaking. The solution adding Triton X-100, tris, and ampholyte (Fig. 2B) was streaked and solution adding Triton X-100, CHAPS, and ampholyte (Fig. 2C) had a spot pattern with the high resolution, but was very expensive. The solution with Triton X-100, and ampholyte was selected as the best 8.0M solutions.

Next, four gels using solutions based on 6.0M urea plus 2.0M thiourea were prepared and the best method was identified. The gels generated with the solution adding Triton X-100, tris, and ampholyte (Fig 3B) and the solution adding Triton X-100, CHAPS, tris, and ampholyte (Fig. 3D) yielded streaked proteins and both had poor

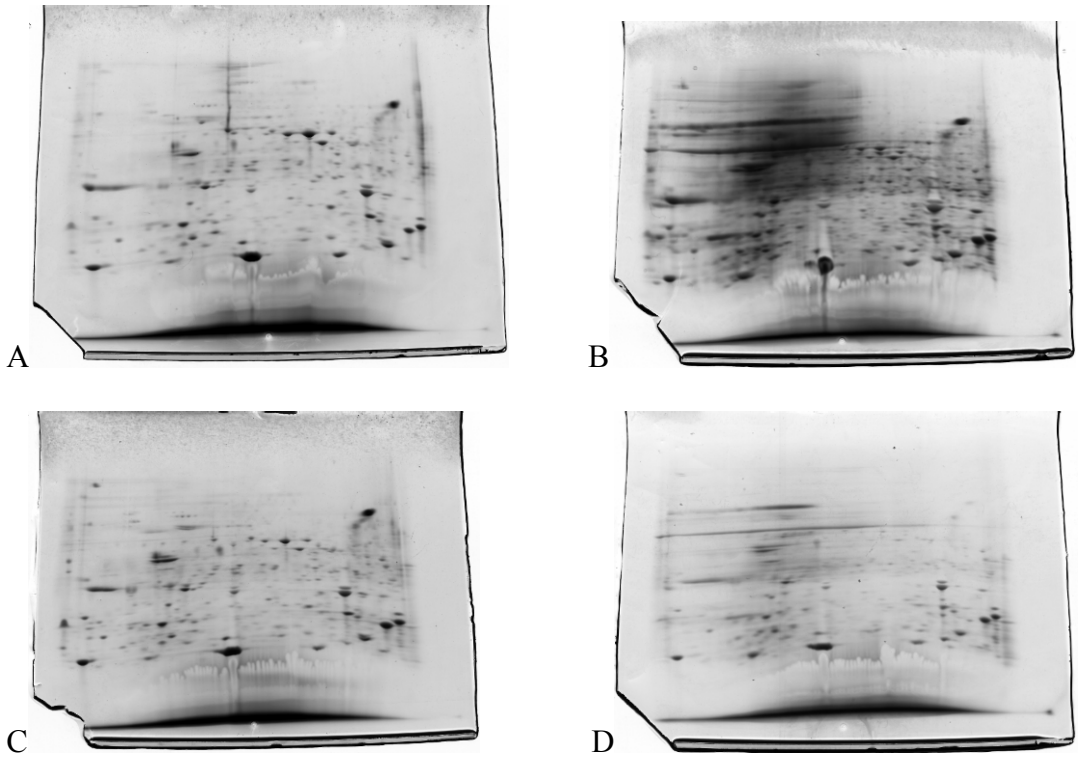


Fig. 3. Examples of gels using 6.0M urea and 2.0M thiourea solutions: (A) 6.0M urea, 2.0M thiourea, Triton X-100, and carrier ampholyte; (B) 6.0M urea, 2.0M thiourea, Triton X-100, tris, and carrier ampholyte; (C) 6.0M urea, 2.0M thiourea, Triton X-100, CHAPS, and carrier ampholyte; and (D) 6.0M urea, 2.0M thiourea, Triton X-100, CHAPS, tris, and carrier ampholyte.

resolution in the higher molecular weight range. The solution adding Triton X-100, and ampholyte (Fig. 3A) and the solution adding Triton X-100, CHAPS, and ampholyte (Fig. 3C) were comparable each having high spot resolution and good clarity. However, due to the relatively high monetary cost of CHAPS, the solution with only Triton X-100 and ampholyte was identified as the best choice from the 6.0M urea plus 2.0M thiourea group, and the best choice overall when compared with the 8.0M urea group.

Part 2: Optimize isoelectric focusing

System development: Optimization of isoelectric focusing — Two-dimensional gels were generated using proteins solubilized using the methods previously described in Objective 1 – Part 1. Three pH ranges were compared: 4–7 pH, 3–10 pH, and 3–10NL pH. Each pH range yielded a usable pattern of protein spots. IPG strips having a 4–7 pH range (Fig. 4A) were limited to the central region of the x-axis and did not display the acidic proteins below pH 4 or basic proteins above pH 7. IPG strips having a 3–10 pH range (Fig. 4B) yielded the same basic pattern with less resolution than the 4–7 pH range across the central region. The relatively wider pH range of the 3–10 pH

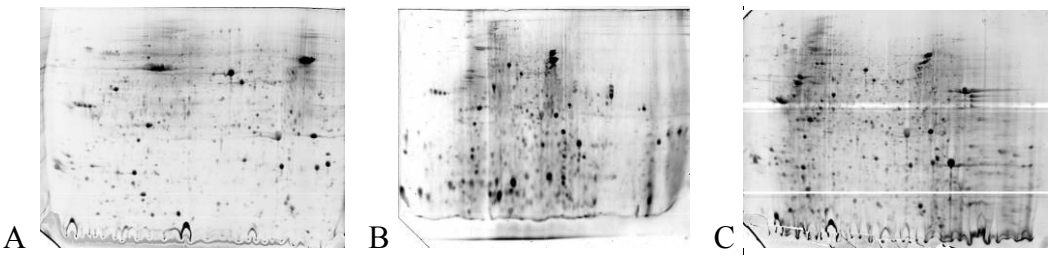


Fig. 4. Examples of gels generated using various pH ranges: (A) 4–7 pH, (B) 3–10 pH, and (C) 3–10NL pH.

IPG strip allowed the display of a greater number of acidic and basic proteins compared with the 4–7 pH IPG strip. Strips having a 3–10NL pH range (Fig. 4C) yielded a pattern

which blended both the 3–10 pH IPG strips and the 4–7 pH IPG strips. These strips yielded a range of proteins similar to the 3–10 pH IPG strips, yet yielded higher resolution in the central region of the gel due to their non-linear pH scale.

The 3–10 pH range was selected as it yielded a large number of proteins distributed along the x-axis while providing a consistent distribution along the range. This distribution also facilitated the arrangements of pI ladders during spot analysis.

Part 3: Optimize SDS-PAGE

System development: Gel fabrication — Two precast gradient gels were tested. These were 8–16% and 10–20% acrylamide gradients. Additionally, 12.0%, 14.0%, 16.0%, and 18.0% acrylamide monomers to move easily through the 12% monomer resulting in vertically compressed spot distribution (Fig. 5A). This monomer percentage may have been useful using shorter run times to compensate for the protein movement. Proteins moved through the 14.0% monomer less quickly, and this system delivered a good spot pattern (Fig. 5B). The 16.0% (Fig. 5C) and 18.0% (Fig. 5D) monomers provided good distributions of protein spots. In the 7cm hand-cast gel system, 18.0% acrylamide monomer out-performed the 16.0% acrylamide monomer. However, in the 24cm system, the 16.0% monomer (Fig. 5E) produced better vertical spot distribution than the 18.0% monomer (Fig. 5F). Gradient gels yielded excellent spot patterns, but their expense made them impractical for use with larger systems.

System development: Optimization of SDS-PAGE gel size — Gels were produced using a variety of systems. These included Ready Gel (7cm), Criterion (11cm), and Ettan DaltSix (24cm) systems. The monetary cost of producing a single gel was directly

proportional to the size of the system. Smaller gels were less expensive and a wide selection of available acrylamide concentrations made them an excellent choice for

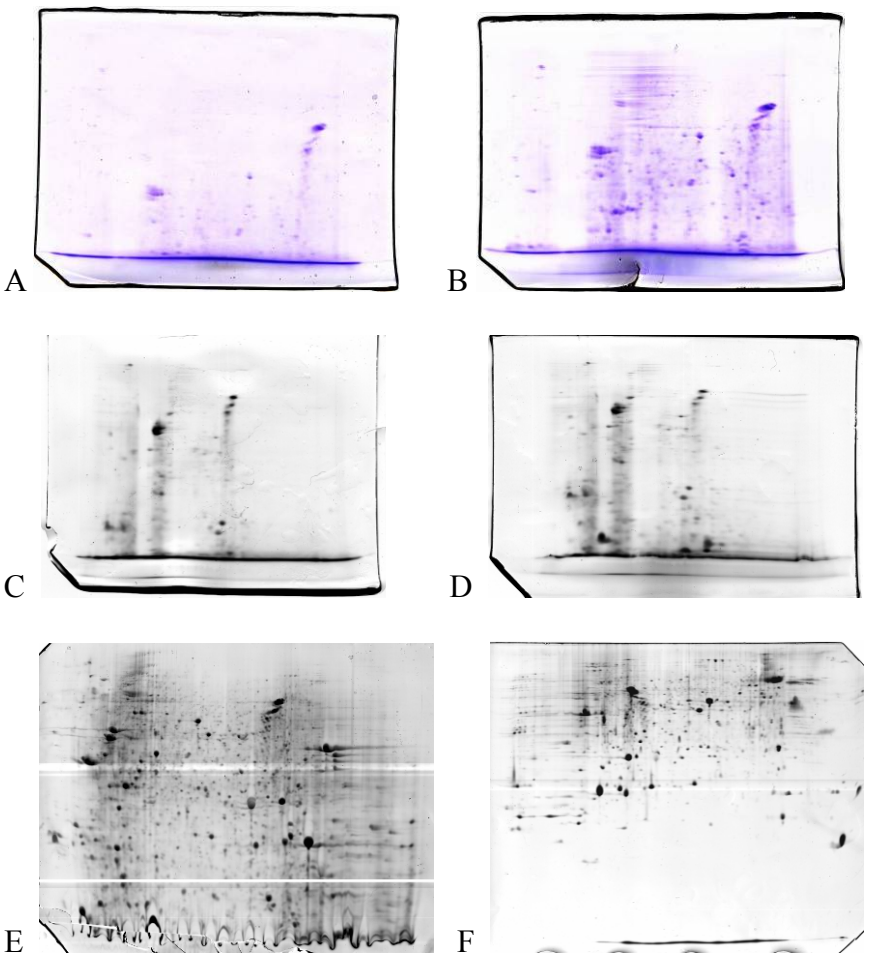


Fig. 5. Differences in spot patterns relative to acrylamide monomer percentage: (A) 7cm with 12.0% monomer; (B) 7cm gel with 14.0% monomer; (C) 7cm gel with 16.0% monomer; (D) 7cm gel with 18.0% monomer; (E) 24cm gel with 16.0% monomer; (F) 24cm gel with 18.0% monomer.

preliminary screening. However, their small size also made them impractical for large scale use (Fig. 6A). The Criterion (11cm) system was tested (Fig. 6B) next and provided good results but had the same disadvantages as the Ready Gel system such as expense relative to hand-cast gels, and less resolution than larger gels. For large scale application,

the Ettan DaltSix system utilizing 24cm gels was chosen (Fig. 6C). The resolution and spot volume of the 24cm system was well suited for spot detection and analysis.

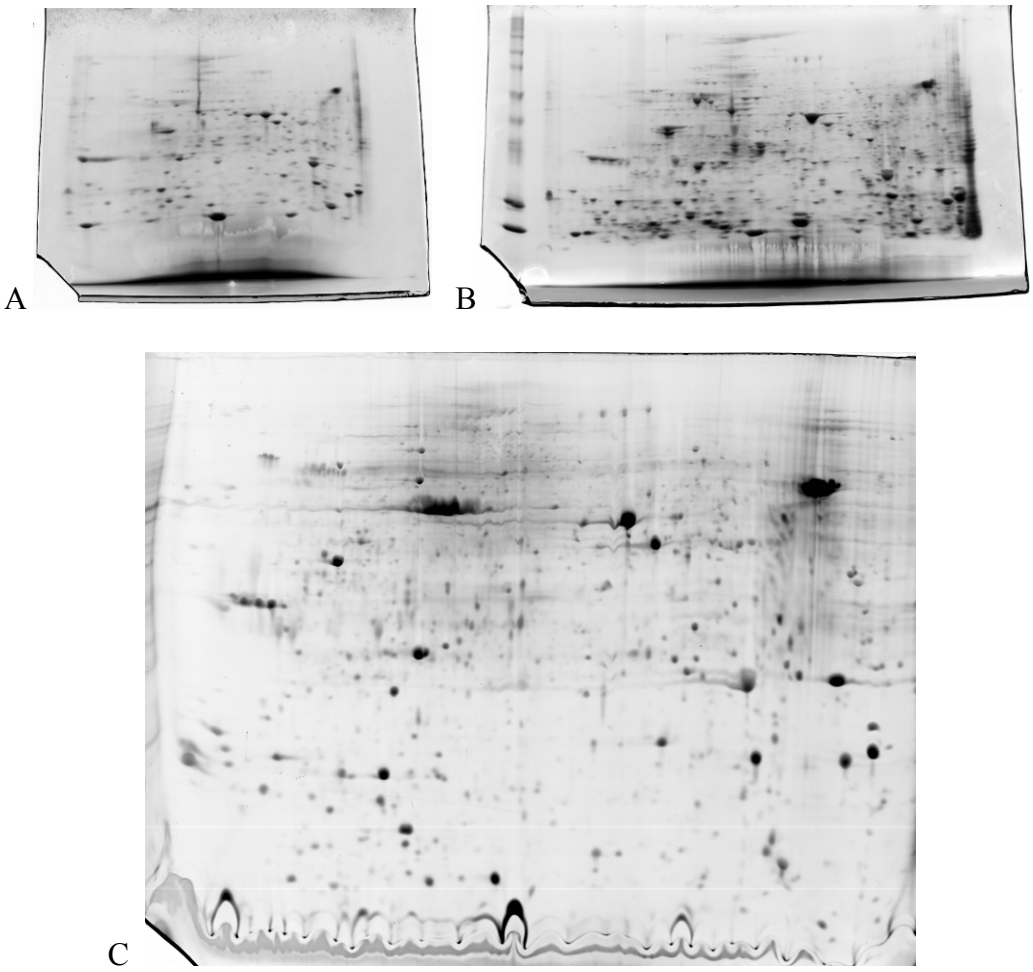


Fig. 6. Comparison of gel sizes: (A) 7cm gel; (B) 11cm gel; and (C) 24cm gel.

System development: Optimization of SDS-PAGE crosslinker — A set of 24cm gels was completed to compare the difference between bis-acrylamide and PDA. The crosslinker was only tested for the 24cm gel size. Due to their large size, the 24cm gels were more fragile and broke easily during handling. PDA is reported to provide a stronger acrylamide gel, reduce streaking in the second dimension, reduce background staining, and decrease the diameter of the protein spots resulting in a better defined spot

pattern (Bio-Rad 1990). PDA was chosen as the crosslinker as it appeared to provide a sharper, more distinct spot. However, significant increase in gel strength was not observed.

System development: Coomassie vs silver stain – CBB R-250 stain was compared with silver nitrate stain. Both stains yielded usable results with each system having unique advantages. Silver stain bound to a greater number of proteins and proteins were detectible at much lower concentrations. This allowed smaller concentrations of protein to be separated and visualized on a gel. However, the smaller quantities of detectable protein were frequently below the minimum amount of protein required to yield acceptable results during mass spectrometry, and would have required collection of protein spots from multiple 2-D gels (Donnelly 2003).

CBB R-250 stain required more protein to yield similar protein detection for visualization. This stain also bound to fewer proteins than silver nitrate and as a result, yielded a similar, but slightly different spot pattern.

Silver nitrate is a hazardous chemical requiring specialized care during handling and disposal. CBB R-250 stain required fewer steps to reach a finalized product. Overall, CBB R-250 was less hazardous to handle and disposal was minimal relative to silver nitrate. Because of the relative safety, while providing acceptable visualization, CBB R-250 was selected as the primary stain.

Objective II – Develop standard protein reference map for *Reticulitermes flavipes*

Using conditions determined to be optimal for separation of termite proteins, a reference map was developed for workers. The Colony 1 24cm gels generated from each replicate were compared and the best gel was selected as the standard reference map. The experimental pI and experimental M_W were estimated using the ladder function in Dymension. Gridlines corresponding to pI and M_W were overlayed on the gel map (Fig. 7) and were useful in visualizing the approximate location of each protein spot. The image (Fig. 8) was processed as described in Materials and Methods, Objective II, Image analysis (pp. 32-33).

As a large number of spots were present on the gel, displaying the Dymension label for each spot on a single figure (Fig. 8) was not feasible. Many labels were obscured by an overlapping label(s). Thus, the labels were removed and the gel map manually processed and numbered, then divided into four quadrants (Fig. 9–12) to facilitate readability. Spots were labeled based on prominence in the gel image. Most spots correspond to consensus spots identified during gel analysis and/or to spots processed for mass spectrometry prior to gel analysis.

Additionally, an overview map and quadrant maps having labeled spots were provided as a digital figure online at http://www.ento.okstate.edu/labs/jwd/index.htm (OSU Entomology and Plant Pathology 2006). Table 3 provides data for each selected spot. The data include spot map number, experimental pI, experimental M_W, and a protein spot identifier. The protein spot identifier was assigned the following format; RF for *Reticulitermes flavipes*, W for worker caste, a four digit pI value, an underscore, and a six digit molecular weight value.

Termites provided an effective subject for two-dimensional gel electrophoresis. Although a whole body extract was used, the proteins separated well even using a wide pH range such as 3–10 pH. Adequate protein was available to facilitate the use of Coomassie brilliant blue staining. With the defined protocols, vertical streaking is rare and horizontal streaking is minimal. A few proteins between 37,000–65,000Da appear to be carbamylated in the 4.0–6.0 pH range. However, very little carbamylation appears to be present based on visual inspection of the termite proteins. A vertical 'streak' of proteins occurs at approximately pH 6.5 beginning at 50,000Da and continuing down until about 20,000Da. This streak is consistant among the gels and appears to be caused by an abundance of lower concentration proteins with a similar pH, but varying M_W. Approximately thirty-six higher concentration protein spots are consistently visible and adequate for use as markers for extrapolating the pI/M_W of other surrounding proteins and for inter-gel correlation of spot patterns. In the 37,000–250,000Da M_W range, there is a large quantity of small spots between 4.0–7.5 pH. Below 37,000Da, smaller spots decrease in frequency and below 20,000Da, few small spots are present. The gel region with the highest population of spots occurs in the 4–7 pH range. Few spots occur in the 3–4 pH range, primarily occurring at 30,000Da and again around 16,000Da, with a single large spot below 10,000Da. As this spot is at the edge of the protein front, it may be comprised of multiple low M_W proteins. From 7–8 pH, spots begin to diminish in frequency as the pH increases. After 8pH, easily visible spots are located mostly in the lower M_W ranges with a few scattered spots occurring above 25,000Da.

Many proteome studies are undertaken using subjects whose genome has been completely sequenced. These include *Arabidopsis thalia*, *Drosophila melanogaster*, and

humans. Some studies look at individual organs or structures. For example, Alonso and Santaren (2006) explored the proteome of the ribosomes of *Drosophila melanogaster*, identifying 52 spots.

Some non-termite 2-D gel systems demonstrate problems which are difficult to overcome. For example, plants contain a protein called ribulose-1,5-bisphosphate carboxylase/oxygenase, or rubisco. Rubisco dominates large portions of the pH range and M_W range, obscuring proteins that lie in its vicinity, and it also visualizes as degradation products or products that have components of the original protein. Hemolymph proteins appear to be among the most commonly published aspects of insect proteomics. Hemolymph protein analysis of wound and immune proteins from *Drosophila melanogaster* yielded gels with visible vertical streaking (Vierstraete et al. 2004a), and carbamylation of proteins appeared to be present in several regions (Vierstraete et al. 2004b). Sharma et al. (2004) compared the effects of carbamate toxicity on brown leafhoppers. Their 2-D gel system used two gel components, 3.5–7.0 pH and 6.0–10.0 pH. The published gels suffered from distortion throughout the acidic pH range. Gels of *Bothrops* snake venom yield three major horizontal regions of spots with much horizontal and vertical streaking (Serrano et al. 2005). Stadler and Hales (2002) observed differences between parasitized and non-parasitized Australian locusts (*Oedaleus australis*). These gels demonstrated some horizontal smearing and streaking and many proteins appeared to exhibit carbamylation. They also had trouble with vertical streaking caused by lipid contamination. Overall, our system described in Objective I yielded excellent results when compared with other systems.

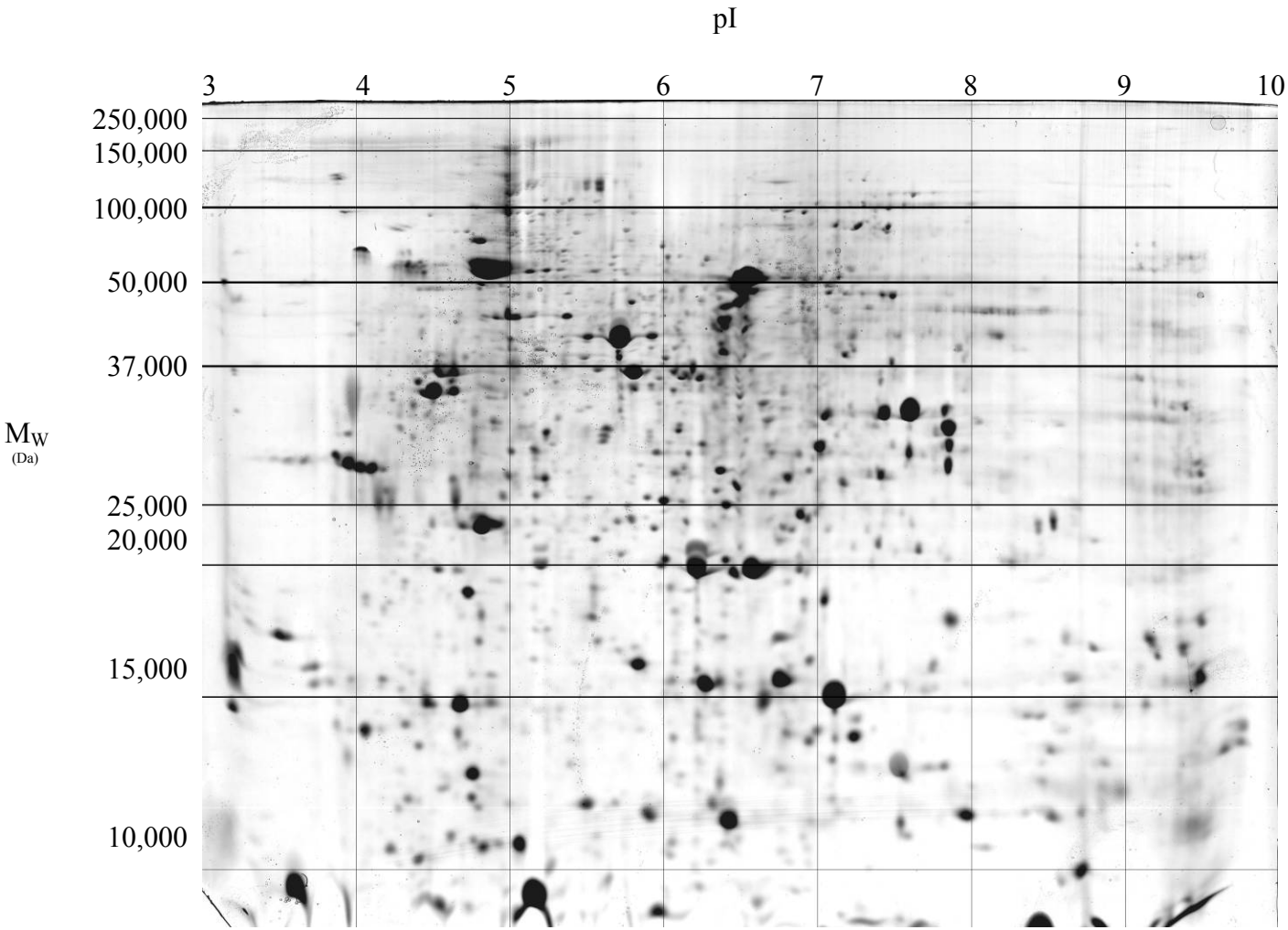


Fig. 7. Grid for estimating pI and molecular weights.

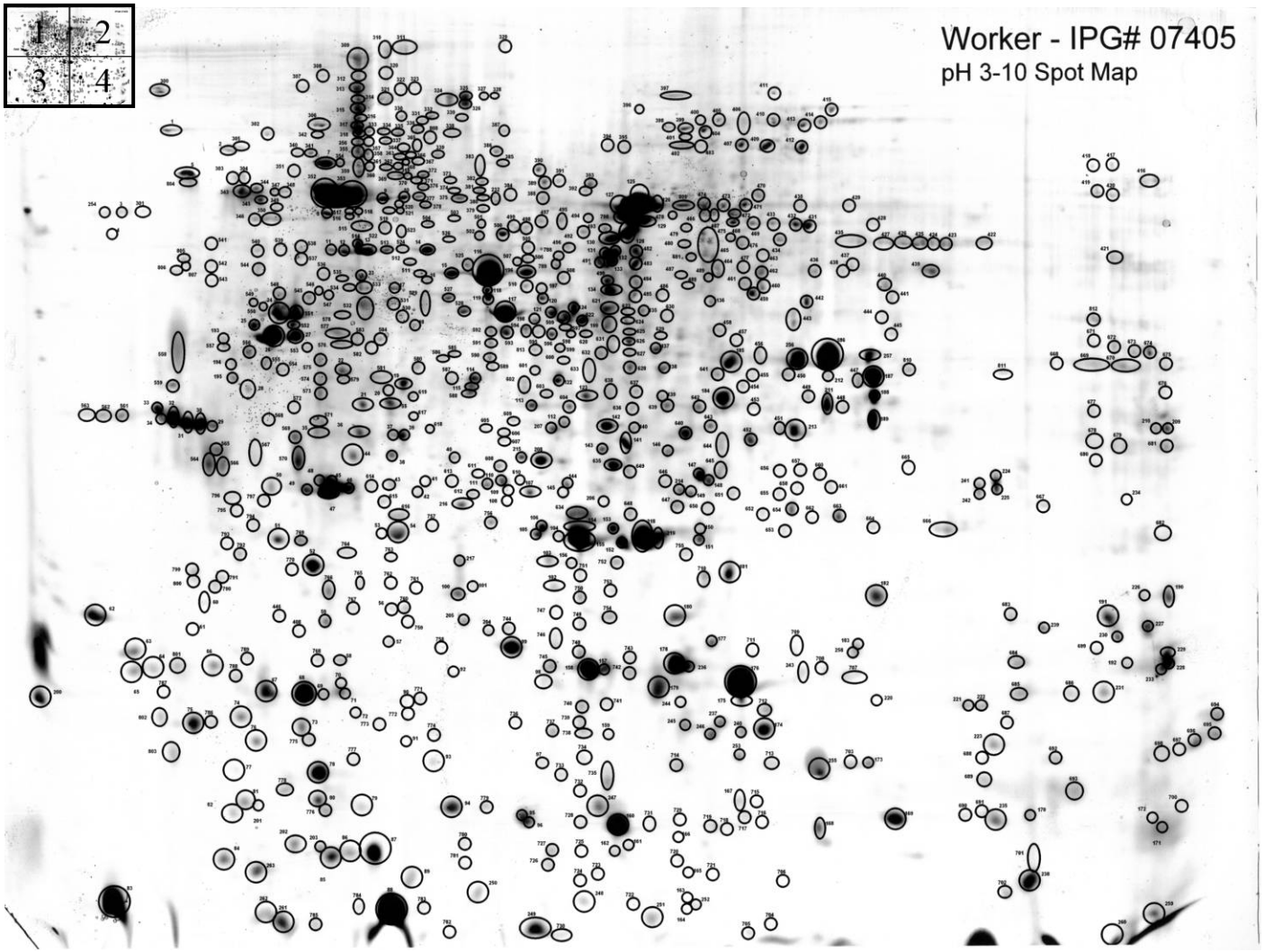


Fig. 8. Worker caste reference Map for *Reticulitermes flavipes*.
Inset: Overview of gel quadrants.

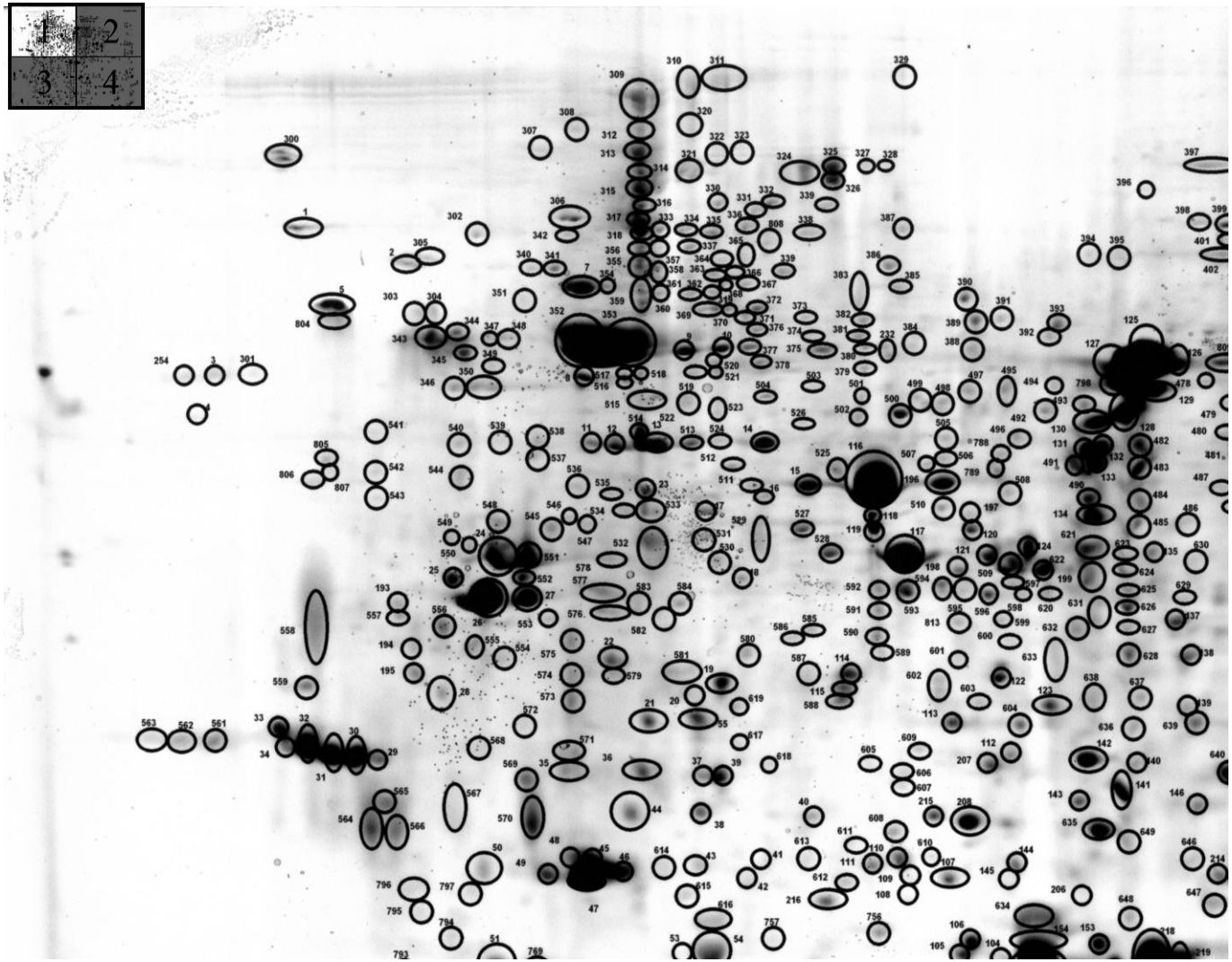


Fig. 9. Worker caste reference map for *Reticulitermes flavipes* – Quadrant 1.
Inset: Highlight of selected gel quadrant.

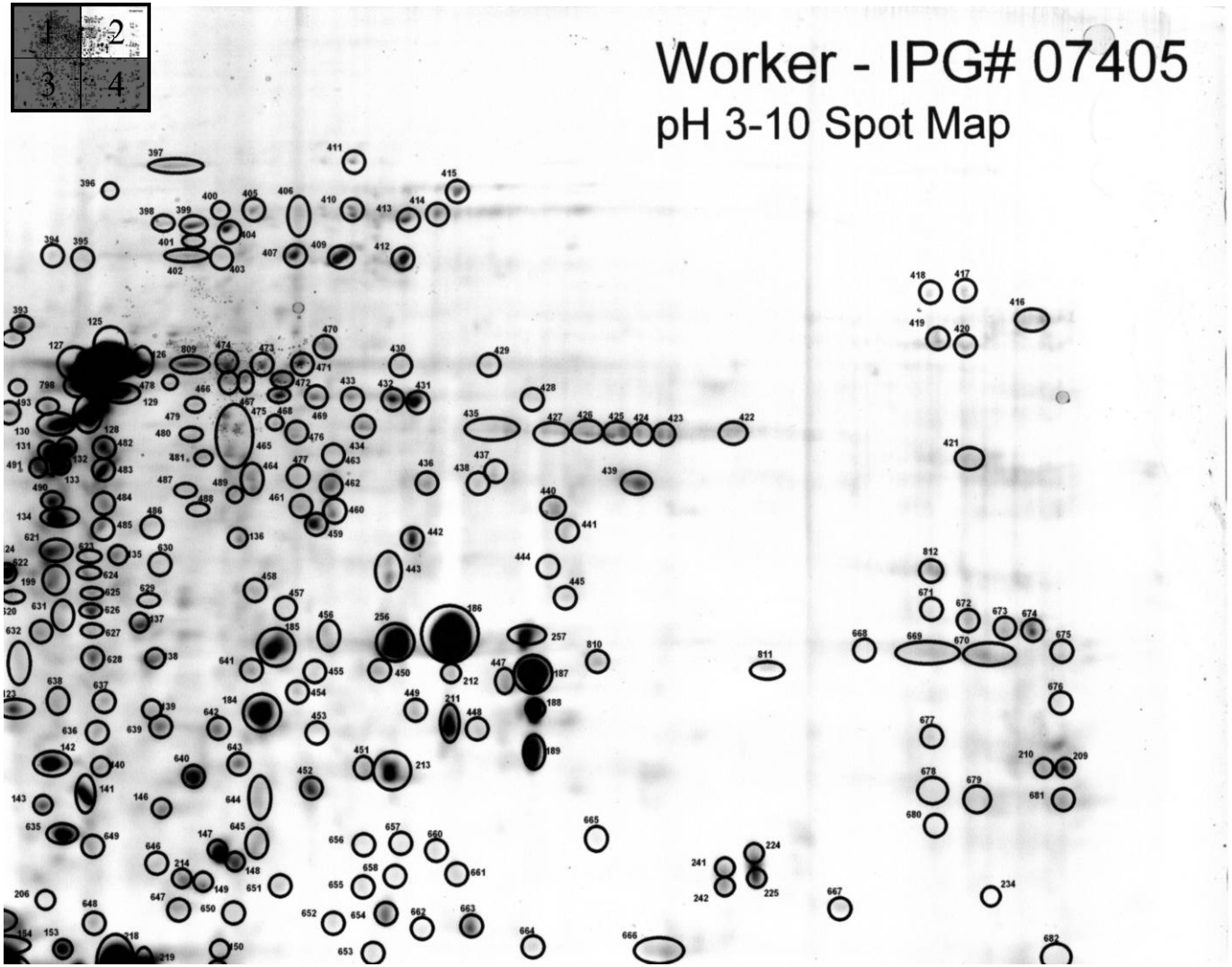


Fig. 10. Worker caste reference map for *Reticulitermes flavipes* – Quadrant 2. Inset: Highlight of selected gel quadrant.

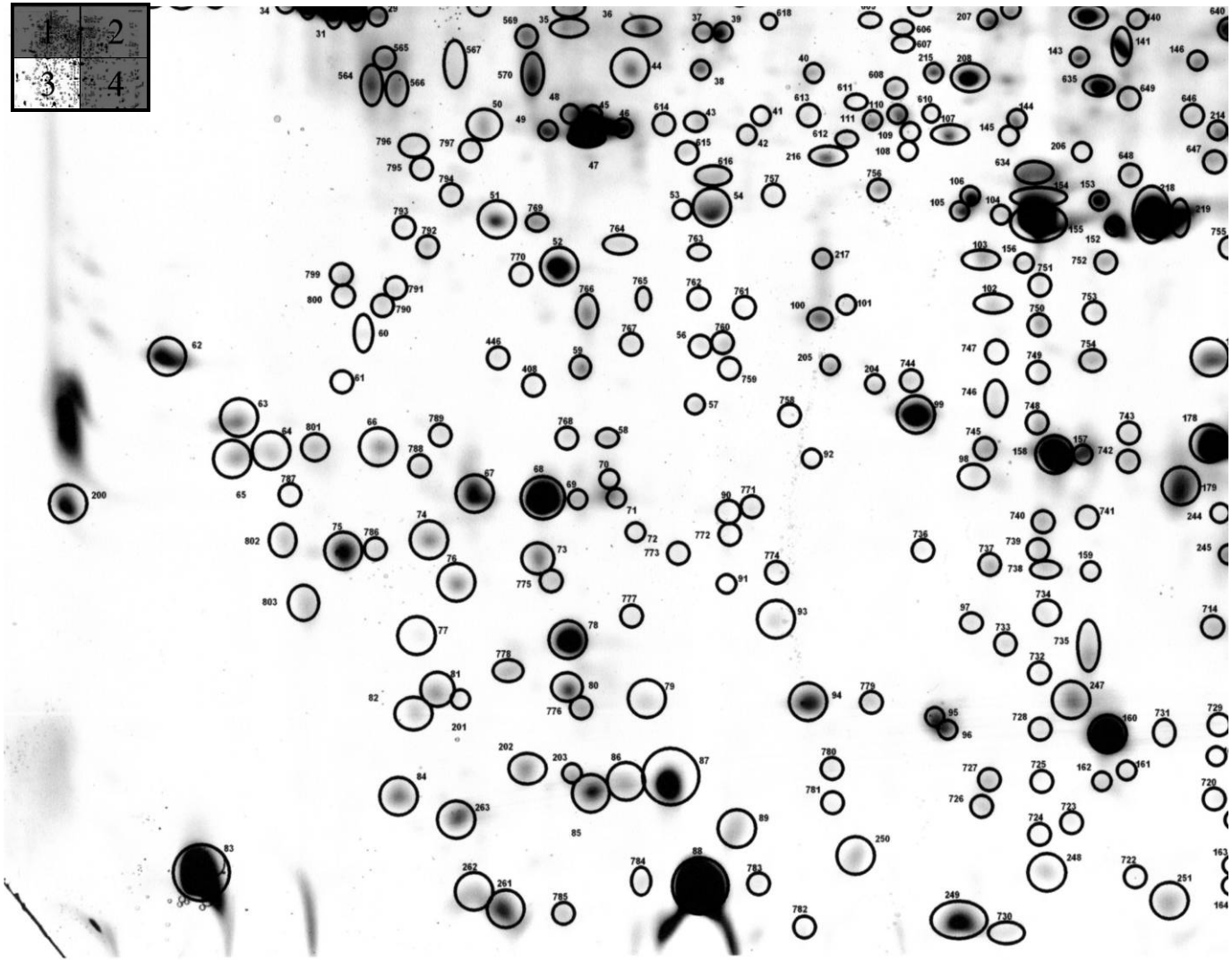


Fig. 11. Worker caste reference map for *Reticulitermes flavipes* – Quadrant 3.
Inset: Highlight of selected gel quadrant.

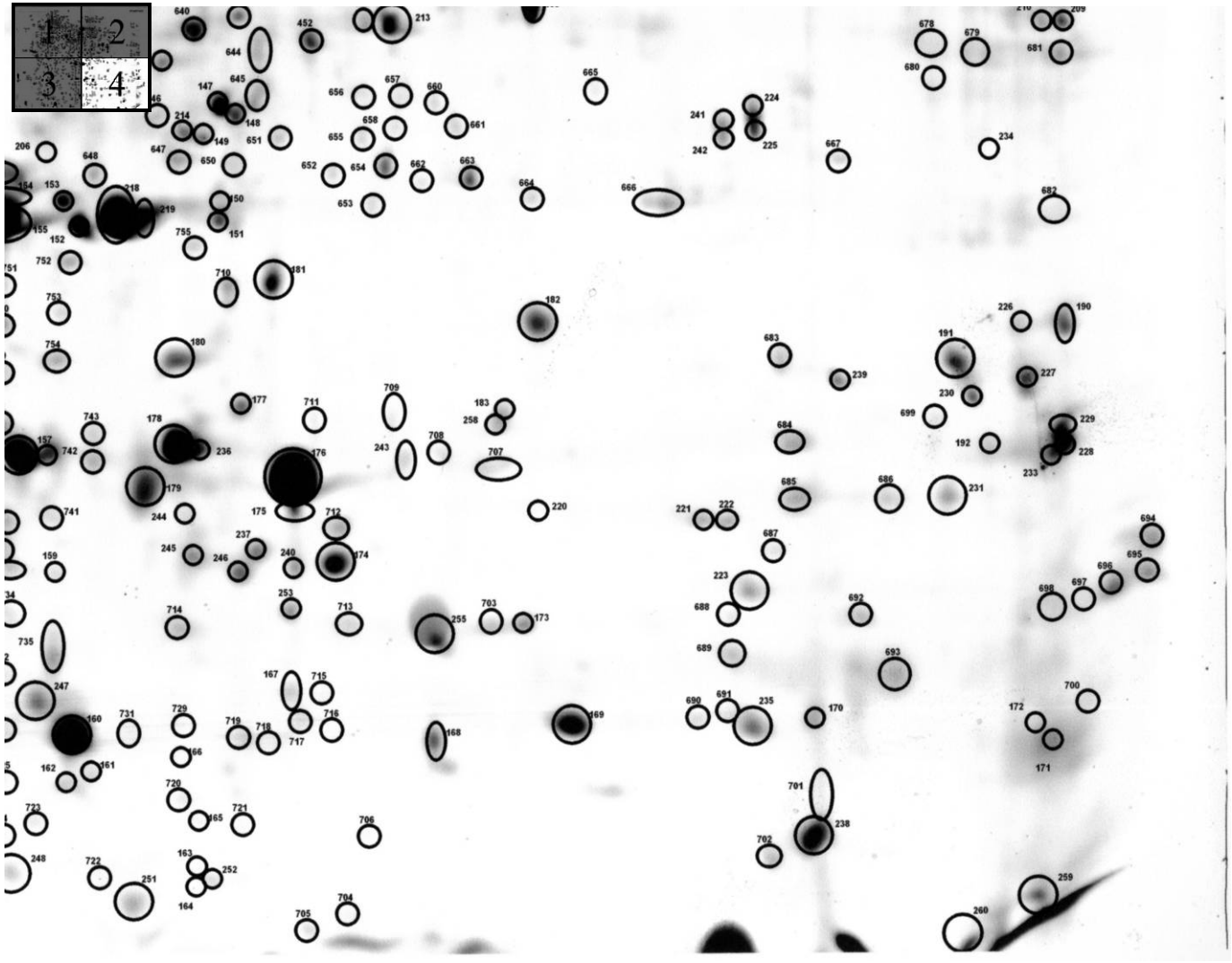


Fig. 12. Worker caste reference map for *Reticulitermes flavipes* – Quadrant 4.
Inset: Highlight of selected gel quadrant.

Table 3. Reference map data

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
1	3.74	89065	RFW0374_089065
2	3.94	89642	RFW0394_089642
3	3.77	72688	RFW0377_072688
4	3.49	67250	RFW0349_067250
5	4.01	67844	RFW0401_067844
6	4.72	100164	RFW0472_100164
7	4.54	97308	RFW0454_097308
8	4.44	81256	RFW0444_081256
9	4.72	79321	RFW0472_079321
10	4.86	77426	RFW0486_077426
11	4.59	53883	RFW0459_053883
12	4.63	54335	RFW0463_054335
13	4.61	56794	RFW0461_056794
14	4.81	64668	RFW0481_064668
15	4.95	56644	RFW0495_056644
16	5.35	57285	RFW0535_057285
17	4.87	47682	RFW0487_047682
18	5.19	41602	RFW0519_041602
19	5.30	34612	RFW0530_034612
20	5.20	34182	RFW0520_064182
21	5.11	31950	RFW0511_031950
22	4.99	35120	RFW0499_035120
23	4.75	49666	RFW0475_049666
24	4.49	51451	RFW0449_051451
25	4.48	39318	RFW0448_039318
26	4.56	38734	RFW0456_038734
27	4.69	39305	RFW0469_039305
28	4.46	33296	RFW0446_033296
29	4.12	34327	RFW0412_034327
30	4.05	29393	RFW0405_029393
31	3.97	29682	RFW0397_039682
32	3.87	29884	RFW0387_029884
33	3.77	29998	RFW0377_029998
34	3.78	35373	RFW0378_035373
35	4.91	28512	RFW0491_028512
36	5.12	29071	RFW0512_029071
37	5.14	33191	RFW0514_033191
38	5.28	27879	RFW0528_027879
39	5.35	29521	RFW0535_029521
40	5.62	26802	RFW0562_026802
41	5.48	25449	RFW0548_025449
42	5.43	24682	RFW0543_024682
43	5.31	24878	RFW0531_024878
44	5.08	26522	RFW0508_026522
45	5.02	24972	RFW0502_024972
46	4.89	27318	RFW0489_027318

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
47	4.94	24245	RFW0494_024245
48	4.73	28240	RFW0473_028240
49	4.65	27120	RFW0465_027120
50	4.69	24164	RFW0469_024164
51	4.47	21387	RFW0447_021387
52	4.68	20266	RFW0468_020266
53	5.04	22913	RFW0504_022913
54	5.37	22535	RFW0537_022535
55	5.28	31923	RFW0528_031923
56	5.14	17834	RFW0514_017834
57	5.43	17673	RFW0543_017676
58	4.87	16732	RFW0487_016732
59	4.74	18141	RFW0474_018141
60	4.03	18234	RFW0403_018234
61	4.00	16674	RFW0400_016674
62	3.41	17625	RFW0341_017625
63	3.63	16285	RFW0363_016285
64	3.75	14409	RFW0375_014409
65	3.64	14302	RFW0364_014302
66	4.10	16080	RFW0410_016080
67	4.41	15485	RFW0441_015485
68	4.67	15450	RFW0467_015450
69	4.75	13008	RFW0475_013008
70	4.84	13764	RFW0484_013764
71	4.88	15811	RFW0488_015811
72	4.95	14976	RFW0495_014976
73	4.61	14460	RFW0461_014460
74	4.25	14593	RFW0425_014593
75	3.98	14401	RFW0398_014401
76	4.37	14137	RFW0437_014137
77	4.22	13412	RFW0422_013412
78	4.70	13319	RFW0470_013319
79	4.99	12602	RFW0499_012602
80	4.72	12704	RFW0472_012704
81	4.29	12749	RFW0429_012749
82	4.20	12431	RFW0420_012431
83	3.58	10573	RFW0358_010573
84	4.15	11457	RFW0415_011457
85	4.79	7216	RFW0479_007216
86	4.88	7467	RFW0488_007467
87	5.00	7326	RFW0500_007326
88	5.16	6220	RFW0516_006220
89	5.24	6889	RFW0524_006889
90	5.52	15871	RFW0552_015871
91	5.57	14940	RFW0557_014940
92	5.73	16747	RFW0573_016747

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
93	5.71	13952	RFW0571_013952
94	5.85	12752	RFW0585_012752
95	5.92	8340	RFW0592_008340
96	5.92	8309	RFW0592_008309
97	6.29	13826	RFW0629_013826
98	6.03	17694	RFW0603_017694
99	5.99	18483	RFW0599_018483
100	5.73	20008	RFW0573_020008
101	5.82	20524	RFW0582_020524
102	6.08	21735	RFW0608_021735
103	6.08	22813	RFW0608_022813
104	6.13	22942	RFW0613_022942
105	5.95	23523	RFW0595_023523
106	6.07	23895	RFW0607_023895
107	6.05	25156	RFW0605_025156
108	5.81	25934	RFW0581_025934
109	5.81	26958	RFW0581_026958
110	5.91	25555	RFW0591_025555
111	5.81	25476	RFW0581_025476
112	5.96	29632	RFW0596_029632
113	5.95	36640	RFW0595_036640
114	5.54	35649	RFW0554_035649
115	5.61	39288	RFW0561_039288
116	5.17	54053	RFW0517_054053
117	5.16	45647	RFW0516_045647
118	5.11	49942	RFW0511_049942
119	5.08	49483	RFW0508_049483
120	5.24	44960	RFW0524_044960
121	6.09	50838	RFW0609_050838
122	5.71	33979	RFW0571_033979
123	5.86	31742	RFW0586_031742
124	5.38	41226	RFW0538_041226
125	6.13	22942	RFW0613_022942
126	5.83	74663	RFW0583_074663
127	6.50	76258	RFW0650_076258
128	6.51	67289	RFW0651_067289
129	5.74	71500	RFW0574_071500
130	5.45	66430	RFW0545_066430
131	5.39	44843	RFW0539_044843
132	6.56	63552	RFW0656_063552
133	6.55	60419	RFW0655_060419
134	5.41	40788	RFW0541_040788
135	5.57	38534	RFW0557_038534
136	6.01	53085	RFW0601_053085
137	5.75	36535	RFW0575_036535
138	5.85	35393	RFW0585_035393

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
139	5.98	33947	RFW0598_033947
140	5.90	31868	RFW0590_031868
141	5.96	30693	RFW0596_030693
142	5.92	30992	RFW0592_030992
143	6.36	31449	RFW0636_031449
144	6.12	26653	RFW0612_026653
145	6.13	26557	RFW0613_026557
146	6.14	30908	RFW0614_030908
147	6.40	29766	RFW0640_029766
148	6.49	29195	RFW0649_029195
149	6.40	28792	RFW0640_028792
150	6.93	24032	RFW0693_024032
151	6.62	25617	RFW0662_025617
152	6.20	25651	RFW0620_025651
153	6.18	26145	RFW0618_026145
154	6.12	25451	RFW0612_025451
155	6.13	24721	RFW0613_024721
156	6.17	20909	RFW0617_020909
157	6.40	18246	RFW0640_018246
158	6.30	18089	RFW0630_018089
159	6.64	14485	RFW0664_014485
160	6.44	8270	RFW0644_008270
161	6.51	7596	RFW0651_007596
162	6.42	7466	RFW0642_007466
163	6.90	6303	RFW0690_006303
164	6.87	6285	RFW0687_006285
165	6.84	7000	RFW0684_007000
166	6.84	8215	RFW0584_008215
167	7.14	9069	RFW0714_009069
168	7.60	8215	RFW0760_008215
169	8.04	8387	RFW0804_009387
170	8.84	8488	RFW0884_008488
171	9.62	8008	RFW0962_008008
172	9.57	8214	RFW0957_008214
173	7.91	10224	RFW0791_010224
174	7.29	11554	RFW0729_011554
175	7.16	12818	RFW0716_012818
176	7.26	16382	RFW0726_016382
177	6.97	18550	RFW0697_018550
178	6.84	17635	RFW0684_017635
179	6.78	16844	RFW0678_016844
180	6.67	20462	RFW0667_020462
181	6.85	23023	RFW0685_023023
182	7.95	18524	RFW0795_018524
183	7.83	15508	RFW0783_015508
184	6.36	33803	RFW0636_033803

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
185	6.30	36043	RFW0630_036043
186	6.97	43594	RFW0697_043594
187	7.34	40554	RFW0734_040554
188	7.36	34211	RFW0736_034211
189	7.43	33059	RFW0743_033059
190	9.63	17770	RFW0963_017770
191	9.31	16871	RFW0931_016871
192	9.35	14594	RFW0935_014594
193	4.30	38056	RFW0430_038056
194	4.22	42676	RFW0422_042676
195	4.38	33683	RFW0438_033683
196	5.30	50084	RFW0530_050084
197	6.02	55139	RFW0602_055139
198	5.98	50376	RFW0598_050376
199	5.61	36563	RFW0561_036563
200	3.42	17442	RFW0342_017442
201	4.34	8866	RFW0434_008866
202	4.53	11862	RFW0453_011862
203	4.72	7518	RFW0472_007518
204	5.95	18590	RFW0595_018590
205	5.80	18695	RFW0580_018695
206	6.13	27540	RFW0613_027540
207	6.00	28929	RFW0600_028929
208	6.09	27416	RFW0609_027416
209	9.71	33764	RFW0971_033764
210	9.63	33718	RFW0963_033718
211	7.09	33382	RFW0709_033382
212	7.02	35105	RFW0702_035105
213	6.93	32140	RFW0693_032140
214	6.81	27015	RFW0681_027015
215	5.99	27069	RFW0599_027069
216	5.68	24570	RFW0568_024570
217	5.74	21583	RFW0574_021583
218	6.24	26767	RFW0624_026767
219	6.71	22934	RFW0671_002934
220	7.93	12907	RFW0793_012907
221	8.45	12627	RFW0845_012627
222	8.55	12625	RFW0855_012625
223	8.63	10918	RFW0863_010918
224	8.45	29940	RFW0845_029940
225	8.65	27104	RFW0865_027104
226	9.48	17693	RFW0948_017693
227	9.54	16679	RFW0954_016679
228	9.61	14518	RFW0961_014518
229	9.63	15030	RFW0963_015030
230	9.36	16435	RFW0936_016435

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
231	9.28	13170	RFW0928_013170
232	5.78	78451	RFW0578_078451
233	9.54	13823	RFW0954_013823
234	9.35	26747	RFW0935_026747
235	8.67	8243	RFW0867_008243
236	6.86	14473	RFW0686_014473
237	7.05	11794	RFW0705_011794
238	8.83	6724	RFW0883_006724
239	8.96	16832	RFW0896_016832
240	7.17	11658	RFW0717_011658
241	8.34	29611	RFW0834_029611
242	8.54	26664	RFW0854_026664
243	7.62	16178	RFW0762_016178
244	6.82	12745	RFW0682_012745
245	6.83	11854	RFW0683_011854
246	7.00	11230	RFW0700_011230
247	6.34	8771	RFW0634_008771
248	6.28	6097	RFW0628_006097
249	5.97	5726	RFW0597_005726
250	5.59	6440	RFW0559_006440
251	6.65	5927	RFW0665_006927
252	7.01	6153	RFW0701_006153
253	7.14	10551	RFW0714_010551
254	3.48	72819	RFW0348_072819
255	7.61	10115	RFW0761_010115
256	6.72	36296	RFW0672_036296
257	7.35	43835	RFW0735_043835
258	7.86	17151	RFW0786_017151
259	9.56	6317	RFW0956_006317
260	9.34	5439	RFW0934_005439
261	4.46	5858	RFW0446_005858
262	4.40	5958	RFW0440_005958
263	4.34	6886	RFW0434_006886
300	3.80	112293	RFW0380_112293
301	3.79	54512	RFW0379_054512
302	4.20	119431	RFW0420_119431
303	3.99	78190	RFW0399_078190
304	4.28	82721	RFW0428_082721
305	4.03	113256	RFW0403_113256
306	4.58	100162	RFW0458_100162
307	4.51	113508	RFW0451_113508
308	4.67	115078	RFW0467_115078
309	4.99	128812	RFW0499_128812
310	5.12	128112	RFW0512_128112
311	5.20	129680	RFW0520_129680
312	4.83	117520	RFW0483_117520

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
313	4.82	113811	RFW0482_113811
314	4.97	109150	RFW0497_109150
315	4.75	105532	RFW0475_105532
316	4.79	101123	RFW0479_101123
317	4.71	99483	RFW0471_099483
318	4.71	95444	RFW0471_095444
319	5.18	85288	RFW0518_085288
320	5.05	116786	RFW0505_116786
321	4.96	112305	RFW0496_112305
322	5.13	114446	RFW0513_114446
323	5.26	113362	RFW0526_113362
324	5.35	107518	RFW0535_107518
325	5.57	110319	RFW0557_110319
326	5.44	106726	RFW0544_106726
327	5.68	110332	RFW0568_110332
328	5.75	110038	RFW0575_110038
329	5.79	129057	RFW0579_129057
330	5.03	102473	RFW0503_102473
331	5.32	102920	RFW0532_102920
332	5.38	103049	RFW0538_103049
333	4.79	97162	RFW0479_097162
334	4.91	99907	RFW0491_099907
335	5.00	97274	RFW0500_097274
336	5.29	97073	RFW0529_097073
337	4.87	93650	RFW0487_093650
338	5.29	96664	RFW0529_096664
339	5.55	102108	RFW0555_102108
340	4.33	106304	RFW0433_106304
341	4.44	103375	RFW0444_103375
342	4.50	107821	RFW0450_107821
343	4.18	65459	RFW0418_065459
344	4.17	68702	RFW0417_068702
345	4.21	67227	RFW0421_067227
346	4.31	59040	RFW0431_059040
347	4.19	71363	RFW0419_071363
348	4.25	85253	RFW0425_085253
349	4.26	65970	RFW0426_065970
350	4.42	72147	RFW0431_061076
351	4.29	102142	RFW0429_102142
352	4.37	91381	RFW0437_091381
353	4.51	86912	RFW0451_086912
354	4.84	87102	RFW0484_087102
355	4.67	98010	RFW0467_098010
356	4.71	94688	RFW0471_094688
357	5.01	94368	RFW0501_094368
358	5.00	91403	RFW0500_091403

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
359	4.94	83341	RFW0494_083341
360	4.73	89320	RFW0473_089320
361	4.83	88024	RFW0483_088024
362	4.88	85175	RFW0488_085175
363	4.91	88198	RFW0491_088198
364	4.97	90900	RFW0497_090900
365	5.28	95136	RFW0528_095136
366	5.24	93051	RFW524_093051
367	5.01	87439	RFW0501_087439
368	4.94	87439	RFW0494_087439
369	5.18	81724	RFW0518_081724
370	4.93	80596	RFW0493_080596
371	4.99	78613	RFW0499_078613
372	5.00	84221	RFW0500_084221
373	5.49	81753	RFW0549_081753
374	5.51	79210	RFW0551_079210
375	5.15	70034	RFW0515_070034
376	5.33	79767	RFW0533_079767
377	4.97	75314	RFW0497_075314
378	5.34	75036	RFW0534_075036
379	5.29	65037	RFW0529_065037
380	5.68	77245	RFW0568_077245
381	5.65	78977	RFW0565_078977
382	5.67	81583	RFW0567_081583
383	5.66	85064	RFW0566_085064
384	5.48	76176	RFW0548_076176
385	5.78	86944	RFW0578_086944
386	5.75	90695	RFW0575_090695
387	5.80	97153	RFW0580_097153
388	6.02	77018	RFW0602_077018
389	6.03	81456	RFW0603_081456
390	6.00	85010	RFW0600_085010
391	6.11	81611	RFW0611_081611
392	5.58	78807	RFW0558_078807
393	6.30	80854	RFW0630_080854
394	5.83	92622	RFW0583_092622
395	6.50	91768	RFW0650_091768
396	6.29	106872	RFW0629_106872
397	6.50	110202	RFW0650_110202
398	6.30	98774	RFW0630_098774
399	6.39	98539	RFW0639_098539
400	6.56	104127	RFW0656_104127
401	6.34	94909	RFW0634_094909
402	6.27	91622	RFW0627_091622
403	6.42	94031	RFW0642_094031
404	6.96	97563	RFW0696_97563

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
405	7.04	101143	RFW0704_101143
406	6.75	98219	RFW0675_098219
407	6.64	93090	RFW0664_093090
408	4.60	16618	RFW0460_016618
409	7.31	92018	RFW0731_092018
410	6.94	101280	RFW0694_101280
411	7.36	111390	RFW0736_0111390
412	7.52	111390	RFW0752_111390
413	7.52	99323	RFW0752_099323
414	7.63	100454	RFW0763_100454
415	7.69	104761	RFW0769_104761
416	9.57	81072	RFW0957_081072
417	9.33	85928	RFW0933_085928
418	9.21	85802	RFW0921_085802
419	9.23	78136	RFW0923_078136
420	9.33	77272	RFW0933_077272
421	9.33	61680	RFW0933_061680
422	8.59	65167	RFW0859_065167
423	8.36	65023	RFW0836_065023
424	8.27	65160	RFW0827_065160
425	8.02	65976	RFW0802_065976
426	7.84	66451	RFW0784_066451
427	7.97	66532	RFW0797_066532
428	7.94	69781	RFW0794_069781
429	7.78	74727	RFW0778_074727
430	6.95	74217	RFW0695_074217
431	7.56	69502	RFW0756_069502
432	6.88	71376	RFW0688_071376
433	7.35	70100	RFW0735_070100
434	7.39	66475	RFW0739_066475
435	7.43	67254	RFW0743_067254
436	7.59	58942	RFW0759_058942
437	7.82	60504	RFW0782_060504
438	7.25	60074	RFW0725_060074
439	8.27	59129	RFW0827_059129
440	7.58	56398	RFW0758_056398
441	8.06	54506	RFW0806_054506
442	6.86	53067	RFW0686_053067
443	7.49	49422	RFW0749_049422
444	7.51	51576	RFW0751_051576
445	8.04	47253	RFW0804_047253
446	4.50	17894	RFW0450_017894
447	7.20	35126	RFW0720_035126
448	7.18	33672	RFW0718_033672
449	6.91	34078	RFW0691_034078
450	6.66	35139	RFW0666_035139

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
451	6.78	32139	RFW0678_032139
452	6.65	31599	RFW0665_034599
453	6.54	33595	RFW0654_033595
454	6.46	34432	RFW0646_034432
455	6.47	35121	RFW0647_035121
456	6.53	36144	RFW0653_036144
457	6.30	37580	RFW0630_037580
458	7.04	48199	RFW0704_048199
459	6.36	55162	RFW0636_055162
460	7.29	55607	RFW0729_055607
461	7.18	56724	RFW0718_056724
462	6.13	60162	RFW0613_060162
463	6.55	62859	RFW0655_062859
464	6.10	57392	RFW0610_057392
465	6.00	67163	RFW0600_067163
466	6.95	72430	RFW0695_072430
467	7.01	72632	RFW0701_072632
468	6.32	69053	RFW0632_069053
469	6.50	70822	RFW0650_070822
470	6.63	77781	RFW0663_077781
471	6.45	74462	RFW0645_074462
472	7.13	72576	RFW0713_072576
473	7.07	74696	RFW0707_074696
474	6.12	75807	RFW0612_075807
475	7.11	66383	RFW0711_066383
476	6.36	65528	RFW0636_065528
477	7.17	59510	RFW0717_059510
478	6.78	72037	RFW0678_072037
479	5.84	68883	RFW0584_068883
480	6.85	69123	RFW0685_069123
481	5.87	63637	RFW0587_063637
482	5.43	63180	RFW0543_063180
483	5.38	41754	RFW0538_041754
484	5.44	40713	RFW0544_040713
485	5.48	39605	RFW0548_039605
486	5.62	39999	RFW0562_039999
487	5.69	61760	RFW0569_061760
488	5.78	58727	RFW0578_058727
489	6.04	58362	RFW0604_058362
490	5.36	41889	RFW0536_041889
491	5.36	44697	RFW0536_044697
492	5.45	67803	RFW0545_067803
493	6.35	69359	RFW0638_068796
494	6.37	72714	RFW0637_072714
495	5.48	71424	RFW0548_071424
496	5.39	49680	RFW0539_049680

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
497	6.01	70878	RFW0601_070878
498	5.42	55865	RFW0542_055865
499	5.84	69881	RFW0584_069881
500	5.37	57577	RFW0537_057577
501	5.67	74115	RFW0567_074115
502	5.25	59396	RFW0525_059396
503	5.10	66296	RFW0510_066296
504	4.91	69456	RFW0491_069456
505	5.38	53927	RFW0538_053927
506	5.34	52029	RFW0534_052029
507	5.88	61244	RFW0588_061244
508	5.30	45384	RFW0530_045384
509	5.40	40427	RFW0540_040427
510	5.26	47595	RFW0526_047595
511	5.30	59308	RFW0530_059308
512	5.26	60843	RFW0526_060843
513	4.63	58841	RFW0463_058841
514	4.52	62800	RFW0452_062800
515	4.61	75153	RFW0461_075153
516	4.91	72480	RFW0491_072480
517	4.53	81488	RFW0453_081488
518	4.56	80201	RFW0456_080201
519	4.69	74056	RFW0469_074056
520	4.83	76211	RFW0483_076211
521	4.83	74046	RFW0483_074046
522	5.10	69109	RFW0510_069109
523	4.75	70962	RFW0475_070962
524	4.70	66323	RFW0470_066323
525	5.08	56862	RFW0508_056862
526	5.02	63067	RFW0502_063067
527	5.47	54162	RFW0547_054162
528	5.11	45118	RFW0511_045118
529	5.04	44824	RFW0504_044824
530	5.05	43805	RFW0505_043805
531	4.98	44667	RFW0498_044667
532	5.01	43171	RFW0501_043171
533	4.88	45814	RFW0488_045814
534	4.83	46144	RFW0483_046144
535	4.74	48775	RFW0474_048775
536	4.74	58675	RFW0474_058675
537	4.60	50544	RFW0460_050544
538	4.54	53759	RFW0454_053759
539	4.56	50700	RFW0456_050700
540	4.42	51269	RFW0442_051269
541	4.12	50559	RFW0412_050559
542	4.20	46706	RFW0420_046706

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
543	4.10	57328	RFW0410_057328
544	4.42	47258	RFW0442_047258
545	4.67	54037	RFW0467_054037
546	4.74	55435	RFW0474_055435
547	4.85	44740	RFW0485_044740
548	4.55	43941	RFW0455_043941
549	6.26	52838	RFW0626_052838
550	4.39	50866	RFW0439_050866
551	4.61	41950	RFW0461_041950
552	4.66	39914	RFW0466_039914
553	4.76	38096	RFW0476_038096
554	4.57	34854	RFW0457_034854
555	4.54	36759	RFW0454_036759
556	4.45	37117	RFW0445_037117
557	4.29	37078	RFW0429_037078
558	3.90	45318	RFW0390_045318
559	3.88	31971	RFW0388_031971
561	3.59	29496	RFW0359_029496
562	3.45	29288	RFW0345_029288
563	3.37	35760	RFW0337_035760
564	4.08	26329	RFW0408_026329
565	4.15	26948	RFW0415_026948
566	4.16	25379	RFW0416_025379
567	4.33	26057	RFW0433_026057
568	4.44	35099	RFW0444_035099
569	4.78	28485	RFW0478_027485
570	4.79	26268	RFW0479_026268
571	4.73	34666	RFW0473_034666
572	4.71	31765	RFW0471_031765
573	4.87	32871	RFW0487_032871
574	4.88	35826	RFW0488_035826
575	4.84	38094	RFW0484_038094
576	4.93	39530	RFW0493_039530
577	4.90	40875	RFW0490_040875
578	4.84	50895	RFW0484_050895
579	4.88	40411	RFW0488_040411
580	5.36	37441	RFW0536_037441
581	5.15	35492	RFW0515_035492
582	5.01	45637	RFW0501_045637
583	5.00	39455	RFW0500_039455
584	5.11	40286	RFW0511_040286
585	5.32	39909	RFW0532_039909
586	5.39	38305	RFW0539_039305
587	5.53	35296	RFW0553_035296
588	5.63	33455	RFW0563_033455
589	5.47	37803	RFW0547_037803

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
590	5.36	40050	RFW0536_040050
591	5.71	45957	RFW0571_045957
592	5.71	47700	RFW0571_047700
593	5.28	42471	RFW0528_042471
594	5.35	41323	RFW0535_041323
595	5.39	40616	RFW0539_040616
596	6.08	47398	RFW0608_047398
597	5.51	38213	RFW0551_038213
598	5.51	38146	RFW0551_038146
599	5.55	37258	RFW0555_037258
600	6.11	45044	RFW0611_045044
601	5.98	41532	RFW0598_041532
602	5.67	34392	RFW0567_034392
603	6.04	38249	RFW0604_038249
604	5.90	30820	RFW0590_030820
605	5.76	30070	RFW0576_030070
606	5.80	33214	RFW0580_033214
607	5.85	28469	RFW0585_028469
608	5.84	26704	RFW0584_026704
609	5.82	31568	RFW0582_31568
610	6.00	25980	RFW0600_025980
611	5.74	26127	RFW0574_026127
612	5.61	26776	RFW0561_026776
613	5.60	25415	RFW0560_025415
614	5.20	24598	RFW0520_024598
615	5.28	23822	RFW0528_023822
616	5.35	23520	RFW0535_023520
617	5.42	31827	RFW0542_031827
618	5.48	29617	RFW0548_029617
619	5.25	37925	RFW0525_037925
620	6.26	48285	RFW0626_048285
621	5.53	38238	RFW0553_038238
622	6.25	49813	RFW0625_049813
623	6.51	53188	RFW0651_053188
624	5.57	37567	RFW0557_037567
625	5.58	37431	RFW0558_037431
626	5.63	36457	RFW0563_036457
627	7.50	29173	RFW0750_029173
628	5.71	35080	RFW0571_035080
629	5.77	37101	RFW0577_037101
630	5.71	38750	RFW0571_038750
631	5.67	35503	RFW0567_035503
632	5.71	34459	RFW0571_034459
633	5.76	33415	RFW0576_033415
634	6.15	26122	RFW0615_026122
635	6.03	29396	RFW0603_029396

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
636	5.82	33034	RFW0582_033034
637	5.78	33745	RFW0578_033745
638	5.81	32868	RFW0581_032868
639	5.99	33082	RFW0599_033082
640	6.20	31750	RFW0620_031750
641	6.23	35396	RFW0623_035396
642	6.19	32963	RFW0619_032963
643	6.32	32089	RFW0632_032089
644	6.47	31525	RFW0647_031525
645	6.58	30076	RFW0658_030076
646	6.72	27842	RFW0672_027842
647	6.32	28359	RFW0632_028359
648	6.18	27299	RFW0618_027299
649	6.51	20946	RFW0651_020946
650	6.59	27964	RFW0659_027964
651	6.65	28945	RFW0665_028945
652	6.92	27980	RFW0692_027980
653	7.13	27555	RFW0713_027555
654	7.15	28309	RFW0715_028309
655	6.96	29337	RFW0696_029337
656	6.94	30298	RFW0694_030298
657	7.07	30212	RFW0707_030212
658	7.10	29382	RFW0710_029382
660	7.24	30077	RFW0724_030077
661	7.32	29321	RFW0732_029321
662	7.28	28023	RFW0728_028023
663	7.46	28054	RFW0746_028054
664	7.73	27639	RFW0773_027639
665	7.80	30625	RFW0780_030625
666	8.24	27795	RFW0824_027795
667	8.84	29319	RFW0884_029319
668	8.98	42723	RFW0898_042723
669	9.26	42417	RFW0926_042417
670	9.45	42067	RFW0945_042067
671	9.26	45305	RFW0926_045305
672	9.38	46508	RFW0938_046508
673	9.51	44165	RFW0951_044165
674	9.59	44777	RFW0959_044777
675	9.68	41883	RFW0968_041883
676	9.69	38667	RFW0969_038667
677	9.25	35633	RFW0925_035633
678	9.17	32321	RFW0917_032321
679	9.46	31977	RFW0946_031977
680	9.21	29537	RFW0921_029537
681	9.69	31468	RFW0969_031468
682	9.72	28152	RFW0972_028152

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
683	8.73	17256	RFW0873_017256
684	8.77	14631	RFW0877_014631
685	8.83	13297	RFW0883_013297
686	9.08	12757	RFW0908_012757
687	8.72	11672	RFW0872_011672
688	8.58	10334	RFW0858_010334
689	8.55	9712	RFW0855_009712
690	8.46	8522	RFW0846_008522
691	8.55	8549	RFW0855_008549
692	8.98	10385	RFW0898_010385
693	9.09	9196	RFW0909_009196
694	9.93	12102	RFW0993_012102
695	9.91	11307	RFW0991_011307
696	6.92	14190	RFW0692_014190
697	9.62	10530	RFW0962_010530
698	9.51	10401	RFW0951_010401
699	9.22	15253	RFW0922_015253
700	9.70	8398	RFW0970_008398
701	8.85	7094	RFW0885_007094
702	8.67	6500	RFW0867_006500
703	7.75	10146	RFW0775_010146
704	7.30	5910	RFW0730_005910
705	7.16	5632	RFW0716_005632
706	7.37	6807	RFW0737_006807
707	7.91	15697	RFW0791_015697
708	7.63	14395	RFW0763_014395
708	7.75	16256	RFW0775_016256
709	7.48	15454	RFW0748_015454
710	6.72	22988	RFW0672_022988
711	7.22	18276	RFW0722_018276
712	7.31	12322	RFW0731_012322
713	7.34	10187	RFW0734_010187
714	6.79	10191	RFW0679_010191
715	7.23	8998	RFW0723_008998
716	7.33	8263	RFW0733_008263
717	7.22	8463	RFW0722_008463
719	6.98	8049	RFW0698_008049
720	6.80	7374	RFW0680_007374
721	6.98	6901	RFW0698_006901
722	6.54	6207	RFW0654_006207
723	6.31	6914	RFW0631_006914
724	6.20	6696	RFW0620_006696
725	6.22	7582	RFW0622_007582
726	6.04	7048	RFW0604_007048
727	6.04	7421	RFW0604_007421
728	6.22	8189	RFW0622_008189

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
729	6.81	8439	RFW0681_008439
730	6.15	5652	RFW0615_005652
731	6.62	8207	RFW0662_008207
732	6.23	9460	RFW0623_009460
733	6.12	9777	RFW0612_009777
734	6.53	13843	RFW0653_013843
735	6.40	9853	RFW0640_009853
736	6.01	15890	RFW0601_015890
737	6.25	15214	RFW0625_015214
738	6.43	15042	RFW0643_015042
739	6.41	15533	RFW0641_015533
740	6.36	16375	RFW0636_016375
741	6.48	16416	RFW0648_016416
742	6.55	17772	RFW0655_017772
743	6.49	18790	RFW0649_018790
744	5.99	19188	RFW0599_019188
745	6.07	18581	RFW0607_018581
746	6.06	19817	RFW0606_019817
747	6.06	20696	RFW0606_020696
748	6.18	19269	RFW0618_019269
749	6.11	21036	RFW0611_021036
750	6.12	22236	RFW0612_022236
751	6.11	22805	RFW0611_022805
752	6.18	24126	RFW0618_024126
753	6.18	23085	RFW0618_023085
754	6.25	21021	RFW0625_021021
755	6.49	25193	RFW0649_025193
756	4.74	19291	RFW0474_019291
756	5.92	23674	RFW0592_023674
757	5.55	22975	RFW0555_022975
758	5.40	15529	RFW0540_015529
759	5.54	18480	RFW0554_018480
760	5.46	19120	RFW0546_019120
761	5.27	19057	RFW0527_019057
762	5.40	20023	RFW0540_020023
763	5.38	21154	RFW0538_021154
764	5.11	21141	RFW0511_021141
765	5.19	19722	RFW0519_019722
766	4.95	19057	RFW0495_019057
767	5.21	18948	RFW0521_018948
768	4.73	16893	RFW0473_016893
769	4.58	21138	RFW0458_021138
770	4.53	19851	RFW0453_019851
771	5.61	15683	RFW0561_015683
772	5.53	15295	RFW0553_015295
773	5.07	14985	RFW0507_014985

Spot Map #	Experimental pI	Experimental M_W (Da)	Protein Identifier
774	5.72	14811	RFW0572_014811
775	4.65	14199	RFW0465_014199
776	4.78	12436	RFW0478_012436
777	4.95	13653	RFW0495_013653
778	4.49	13182	RFW0449_013182
779	9.48	8823	RFW0948_008823
780	5.56	7703	RFW0556_007703
781	5.54	7023	RFW0554_007023
782	5.46	5618	RFW0516_005618
783	5.33	6212	RFW0533_006212
784	4.91	6037	RFW0491_006037
785	4.69	5748	RFW0469_005748
786	4.08	14331	RFW0408_014331
787	3.82	15487	RFW0382_015487
788	4.22	15729	RFW0422_015729
789	4.30	15106	RFW0430_015106
790	4.11	19281	RFW0411_019281
791	4.15	19892	RFW0415_019892
792	4.26	21523	RFW0426_021523
793	4.17	22523	RFW0417_022523
794	4.33	22026	RFW0433_022026
795	4.22	22326	RFW0422_022326
796	4.20	23609	RFW0420_023609
797	4.34	23389	RFW6434_023389
798	6.48	73947	RFW0648_073947
799	3.99	20451	RFW0399_020451
800	5.65	35360	RFW0565_035360
801	3.89	14537	RFW0389_014537
802	3.80	12075	RFW0380_012075
803	3.88	13882	RFW0388_013882
804	3.96	82270	RFW0396_082270
805	4.07	46526	RFW0407_046526
807	4.08	45002	RFW0408_045002
808	5.36	96672	RFW0536_096672
808	5.37	96442	RFW0537_096442
809	5.98	75042	RFW0598_075042
810	8.14	41720	RFW0814_041720
811	8.69	41062	RFW0869_041062
812	5.48	39057	RFW0548_039057
813	5.47	38612	RFW0547_038612

End of Table

Objective III – Initial characterization of the putative *R. flavipes* proteome

Initial spots for mass spectrometry were chosen based on relative position and potential reproducibility. Selected spots were assigned a numeric identifier and marked on the gel map (Fig. 13). The mass spectrometry gel map was divided into four quadrants (Fig. 14–17) to facilitate viewing. Since mass spectrometry was begun prior to acquiring Dymension software, several 'non-consensus' protein spots were selected as samples for mass spectrometry.

Putative protein identifications were derived using PMF, CAF, and/or MS/MS. Initial characterization of the *R. flavipes* proteome was undertaken using MALDI-TOF MS, and the resulting peptide mass fingerprints were used to generate peak lists to search protein databases such as MSDB, NCBInr and Uniref. Examples of PMFs yielding high confidence putative identifications are shown in Fig. 18. These PMFs demonstrate qualities desired in MS spectra, distinctive peaks, few external contaminates such as keratin and trypsin, and low 'grass' (short peaks of irrelevant data which can obscure taller more relevant peaks). Manually collected peak lists consisted of 25–50 peaks, selecting the highest peaks over the entire spectra. Most peak lists were assembled using Mascot Daemon (Matrix Science Inc., Boston, MA, USA). Automated peak lists were generated by selecting the five highest peaks for every 100Da of spectra. These peak lists consisted of approximately 150 peaks. Most of the PMFs generated were of excellent quality, yet relatively few yielded identifications of suitable confidence. Many of the spots were selected from the largest spots, and many more spots were selected from spots of intermediate density, while a few peaks yielded no PMF. Only a few spots, selected from among the least visible, failed to yield a PMF. Comparing the physical spot

characteristics of proteins that yielded high confidence putative identifications, there was no obvious trend among spot size, density, or gel location. Each database search (Matrix Science Ltd. 2007) typically provided multiple putative identifications (Fig. 19). The first page was comprised of four basic sections and provided basic information for multiple putative identifications including the protein accession number, the protein name, and the protein source. The first section included general information such as user name, email address, data file, database, and top score. The second section included the ion score information and the "probability based MOWSE score" (**mo**lecular **w**eight **se**arch) bar chart graphing one or more of the top scores (depending on the separation among scores). The third section was comprised of the "concise protein summary report" and included the top queries ranked in descending order from highest score. Each of these identified a 'linked' accession number, the MOWSE score, the expectation score, the number of queries matched, and the corresponding protein name. The MOWSE and expectation scores provided the initial indication of reasonable putative protein identification(s). The fourth section outlined the search parameters and included information such as 'type of search', enzyme used, and modifications. Clicking an accession number link provided a second page specific to the protein selected (Fig. 20). This page provided information including peptide coverage of the protein sequence, missed cleavages required to match the sequence, and a scatter plot of the peptide matches. (Each of these sections provided additional indication to the validity of the putative identification.) Most of the database searches yielded multiple putative identifications for each spot and a corresponding probability based MOWSE score. Many of the MOWSE scores (Pappin et al. 1993) were not statistically significant, implying the

identifications had a higher chance of being due to random peptide matches. Because the *R. flavipes* genome is not sequenced, *R. flavipes* proteins are rare in the protein databases and the putative protein identification was typically to a protein from another species (cross-species identification) having a similar peptide sequence. Proteins matched from other species may be homologous to *R. flavipes* proteins. Homologous proteins may not have the same pI or molecular weight. As a result of this, we cannot assume proteins with differing pI/MWs are not functionally equivalent, since the protein function may be similar.

High confidence putative identifications using PMF were not as numerous as preferred due to the lack of termite genome sequencing and the lack of related information available in the protein databases. To ascertain additional high confidence putative identifications, other methods of mass spectrometry were used. CAF was attempted since it could be performed using the MALDI-TOF mass spectrometer available at the BMB core facility. A group of 96 samples were selected (Fig. 21) and submitted for CAF. Although initial tests were promising and yielded several putative identifications, the method proved to be extremely time-intensive due to two factors. First, it required the generation of a PMF to identify the peptide peaks to be selected for further evaluation (Fig. 22A). Then the sample was sulfonated to make the peptide labile and the resulting peptide fragments could be used to elucidate amino acid sequences of the selected peptide (Fig. 22B). Second, every amino acid sequence had to be manually elucidated from the fragmentation pattern. Once the amino acid sequence was elucidated, it could then be searched against the protein databases for a putative identification. Two trials were undertaken. The first trial was six samples and yielded five PMF and one

CAF spectra. The CAF yielded a cuticular protein putative identification. The second trial was five samples and yielded five PMF and no CAF spectra. These spectra yielded three moderate and two low putative identifications. While a promising technique, the results were below expectations.

The final method employed was tandem mass spectrometry (MS/MS) using an ABI 4700 MALDI-TOF/TOF instrument. MS/MS is similar in concept to CAF in that a protein ion is selected then further broken down. MS/MS can provide higher sensitivity and generate spectra from smaller sample amounts. Spot samples matching the CAF samples (Fig. 23) were prepared and sent to the Nevada Proteomics Center located at the University of Nevada, Reno, NV. This method yielded the highest percentage of high confidence putative identifications, but the high cost (~ $12/spot) made it too expensive for extensive use. Data were available using the internet through Bio-Rad IDQuest software (Fig. 24). MS/MS did provide substantially more high-confidence putative identifications with 33 out of 96 samples for 34.4% compared to PMF with 61 out of 890 samples for 6.9%. MS/MS was also the only method used that provided 100% success in generating peak list data for each sample.

A table of putative identifications was assembled and duplicate identifications of the same spot from different gels were omitted. Of proteins having multiple identifications, putative identifications with the highest scores were included. Each of the putative identifications included a description of the protein and the source against which the protein was compared. Only sixteen putative identifications were from termite sources (Table 4). Interestingly, thirteen of the sixteen 'termite' identifications were for hexamerins which belong to a group of high weight hemolymph proteins that include

hemocyanins, prophenoloxidases, and storage proteins. Hexamerins are comprised of six subunits that can all be the same or may be comprised of up to three different peptide subunits. Subunits typically range from 70 to 85kDa with the holoprotein size being around 500kDa (Telfer and Kunkel 1991, Burmester and Scheller 1999). It is not uncommon for hemolymph to have more than one type of hexamerin. Each hexamerin may fulfill different roles in the insect (Moreira et al. 2004), such as hemolymph storage proteins. For example, Moreira et al. observed two hexamerins, Hex-L and Hex-F, in *Musca domestica* (house fly). Moreira et al. elucidated Hex-L was synthesized in fat bodies and speculated this protein is important in metamorphosis. It was also elucidated Hex-F was synthesized in insect fat bodies following protein meals and speculated to be vital as a source of amino acids for egg formation in females. Danty et al. (1998) found three hexamerins, Hex70a, Hex70b, and Hex70c, to be present in the honeybee antennae cuticle.

Social insects, such as honey bees and ants, are holometabolous. Hexamerins are common in holometabolous insects and the hemolymph proteins will comprise a majority of the protein concentration prior to metamorphosis (Scheller et al. 1990, Telfer and Kunkel 1991). Since insects cannot eat while undergoing metamorphosis, they must have adequate protein stores. Hexamerins may also be present in the adult insect. In honey bees, multiple subunits have been identified in the developing stages, most approximately 70kDa (Hex70a, Hex70b, and Hex70c). In the adult honey bee, fewer hexamerins are present, but 70, 80, and 110kDa subunits can be found in the adult queen (Danty et al. 1998).

Hexamerins are not restricted to holometabolous insects. Lewis et al. (2002) demonstrated a correlation between the hexamerin AgSP-1 and diapause, as well as an inverse correlation between the hexamerin and vitellogenin in the boll weevil (*Anthonomus grandis*). Hexamerins are also found in the only known paurometabolous social insect, termites. Termites do not undergo metamorphosis, so hexamerins must play another role in their development. Recently, it was demonstrated that hexamerin genes *Hex-1* and *Hex-2* are present in *R. flavipes* and appear to be important in the developmental physiology of the termite (Scharf et al. 2005b, Scharf et al. 2005a, Zhou et al. 2006). Zhou et al. (2006) speculate hexamerins may prevent worker caste termites from molting into soldier caste termites by absorbing juvenile hormone (JH). Hexamerins appear to function as JH binding proteins in several insects such as the migratory locust (*Locusta migratoria*), the grasshopper (*Melanoplus sanguinipes*), and termites (*Reticulitermes flavipes*) (Braun and Wyatt 1996, Gilbert et al. 2000, Zhou et al. 2006).

A hexamerin in its native state would be too large to transfer from an IPG strip into an acrylamide gel. Only denatured hexamerins would be visualized. As the native hexamerin denatures, it could result in related peptide variants consisting of different subunits. Degradation products may also be present. These products are generated when larger molecules are converted to smaller ones. In proteins, degradation occurs by the loss of peptide fragments. Fragments may consist of functional subunits, peptides of various length, or combinations of subunits and peptides. Thus, each degrading protein could result in multiple degradation products.

It is also possible for a protein to be present in such quantity to mask a smaller protein with a similar pI and MW. For example, ribulose-1,5-bisphosphate carboxylase/ oxygenase (rubisco) is a major protein in plants. This protein is so predominant it masks other proteins having similar pI/MWs. Rubisco also demonstrates degradation products appearing in multiple locations on a two-dimensional gel. This means it would be possible for larger spots to mask smaller spots so several spots could be present in an area. This is increasingly true for wide pH ranges such as 3–10 pH. Some of the more prominent spots could be multiple proteins overlapping due to inadequate separation. This does not prevent high confidence putative identifications, but can complicate the process.

Of the non-termite sources, most were from other insects or non-insect arthropods such as *Ixodes scapularis* (deer tick or black-legged tick) or *Artemia* sp. (brine shrimp). Examples of high confidence putative identifications include tropomyosin from *Periplaneta americana*, hypothetical protein from *Drosophila melanogaster*, and heat shock protein 20.6 from *Locusta migratoria*. Table 5 contains 310 putative protein identifications and includes data such as the confidence, map spot number, experimental pI, experimental MW, the protein identifier, score, expectation, description, and protein source.

Each of the identifications was indicated with one of four confidence categories: high, good, moderate, or low. High confidence scores were statistically significant (P>0.05) MOWSE scores. Mascot MOWSE scores are based on the analysis algorithms described by Pappin et al. (1993). Mascot adds consideration of probability to the MOWSE score and defines the Mascot algorithm as "$-10*LOG_{10}(P)$ where P is the absolute probability"

(Perkins et al. 1999). This establishes an inverse relationship between the probability and the MOWSE score. Thus, the smaller the probability of the number being due to randomness, the higher the MOWSE score. The significance level was based on a combination of the probability of a random match and the size of the sequence database searched and displayed as a bar graph with non-significant scores being displayed in a green region. MOWSE scores were used to establish high confidence matches and significance levels were used for additional confirmation. However, expectation value ('expect') was also used to evaluate matches approaching high confidence. This value indicated the likelihood of random matches and correlated with the significance threshold. Good confidence scores were just below high confidence, but had an 'expectation value' of ≤ 0.5. Moderate confidence scores were also approaching high confidence, but the 'expect' was between 0.5 and 1.0. All other identifications were considered to be low confidence. Approximately 18.7% (58 of 310) of the putative identifications were considered high-confidence identifications. Another 12.6 % (39 of 310) of the putative identifications were considered good confidence identifications. Additionally, 11.3% (35 of 310) of the putative identifications were considered moderate confidence identifications. The remaining 178 putative identifications were considered low confidence identifications. While the low incidence of acceptable putative identifications was most likely due to the lack of a termite genome, it is also possible the extreme genetic diversity among the various insect orders contributes.

A detailed table of mass spectrometry data can be found on the website (http://www.ento.okstate.edu/labs/jwd/index.htm) or on the attached CD (bound edition). The archived data included the gel map number, putative protein identity, protein spectra,

peak list, theoretical and experimental pI, theoretical and experimental MW, database search score, number of matched peptide fragments/matched fragments, percent coverage, and date of database search. New protein data are submitted to protein databases daily. These submissions may include new proteins, revised protein descriptions, additional sequence data, or even confirmation of theoretical proteins. As the databases improve, the quality of putative identifications of termite proteins should also improve. Archived data will be reevaluated periodically.

Putative protein identifications for the worker caste and soldier caste (also see Objective V) were combined for analysis of molecular function. Removing redundant proteins between castes yielded 440 putative identifications for consideration. Resulting identifications were matched to analogous proteins with known gene sequences, yielding 232 protein matches and corresponding sequences. There were 208 proteins not matched. Percentages of matched and not matched proteins are shown in Fig. 25. Gene sequences were then searched to generate analogous protein functions.

Protein functions were inferred by matching UniProt records to ontologies (controlled vocabularies). The Gene Ontology (GO) Consortium (1999-2007) uses databases from three well-defined model organisms: *Drosophila* sp., mouse, and *Sarccharomyces* sp. Further, three ontologies are provided by the GO consortium that include molecular function, cellular component, and biological process. Each ontology consists of 'terms' relating to molecular function. However, there was no obvious preference in the literature for a particular ontology. For example, a comparison of *Drosophila melanogaster* embryos and adult heads (Tarazka et al. 2005) used 'cellular components' to categorize proteins, while a proteomic study of honey bee sperm

categorized proteins by ‘biological process’. Molecular function and biological process were observed by Cristino et al. (2006).

Termite data were classified using the molecular function ontology terms. Protein functions were assigned to first-level terms: binding, catalytic activity, structural molecule activity, transporter activity, motor activity, signal transducer activity, transcription regulator activity, enzyme regulator activity, antioxidant activity, and translation regulator activity. Many proteins yielded multiple molecular functions with 232 proteins being matched to 329 functions. For example, vacuolar ATP synthase subunit H was linked to four different molecular functions: ‘transporter activity’, ‘enzyme regulator activity’, ‘catalytic activity’, ‘binding activity’. Other proteins were matched to a single function. For example, hexamerins I and II were assigned into ‘transporter activity’. Assigned functions comprised sixty-one percent of the total proteins observed and the remaining thirty-eight percent comprised of proteins with unknown function (Fig. 26). Binding and catalytic activity were the two largest categories with 26% and 22%, respectively. These two categories were separated into sub-categories.

Proteins from the Binding category were assigned to: nucleotide, ion, nucleic acid, protein, tetrapyrrole, cofactor, odorant binding, amine, carbohydrate, vitamin, lipid, steroid, chromatin, and pattern binding (Fig. 27).

Proteins from ‘catalytic activity’ were assigned to: hydrolase, transferase, oxidoreductase, helicase, lyase, isomerase, ligase, cyclase, integrase, and small protein activiating enzyme activity (Fig. 28).

Bevan et al. (1998) established a classification for *Arabidopis thalia* metabolic functions. This classification is common in plant proteomic studies such as Watson et al. (2003) and Donnelly et al. (2005). However, in animals, the Gene Ontology classification appears to be the accepted standard. For example, one study compared the adult head to whole embryos of *Drosophila melanogaster* and found 1,133 total proteins with 307 common proteins between the stages (Tarazka et al. 2005). Tarazka et al. (2005) compared identified proteins to GO cellular functions. Mitochondrion, nucleus, and cytoplasm were the top three protein functions elucidated.

Cristino et al. (2006) observed 'molecular function' with 'nucleic acid binding' and 'structural constituent of ribosome' comprising over fifty percent of the assigned GO terms, followed by protein binding at twelve percent. Interestingly, Cristino et al. (2006) also noted ten EST matches for a hexamerin gene in honey bees indicating hexamerins as a highly abundant protein. Comparing termite protein functions to those of honey bees, the GO term assignments were similar but were expressed in substantially different levels. For example, termite proteins were most common for the 'binding' and 'catalytic activity' GO terms, while honey bee proteins were most common for the 'nucleic acid binding' and 'nucleotide binding' GO terms.

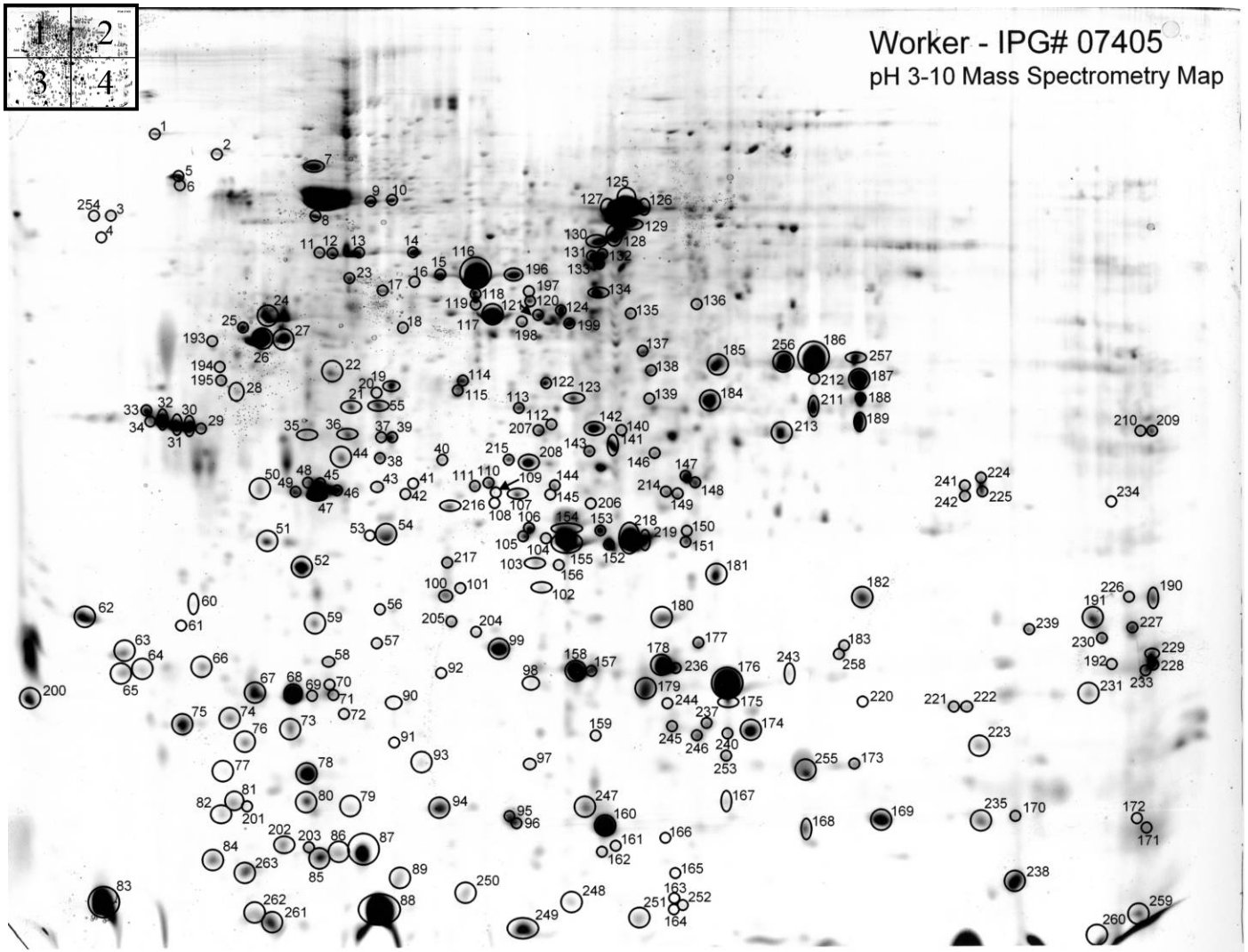


Fig. 13. Worker caste mass spectrometry reference map for *Reticulitermes flavipes*.
Inset: Overview of gel quadrants.

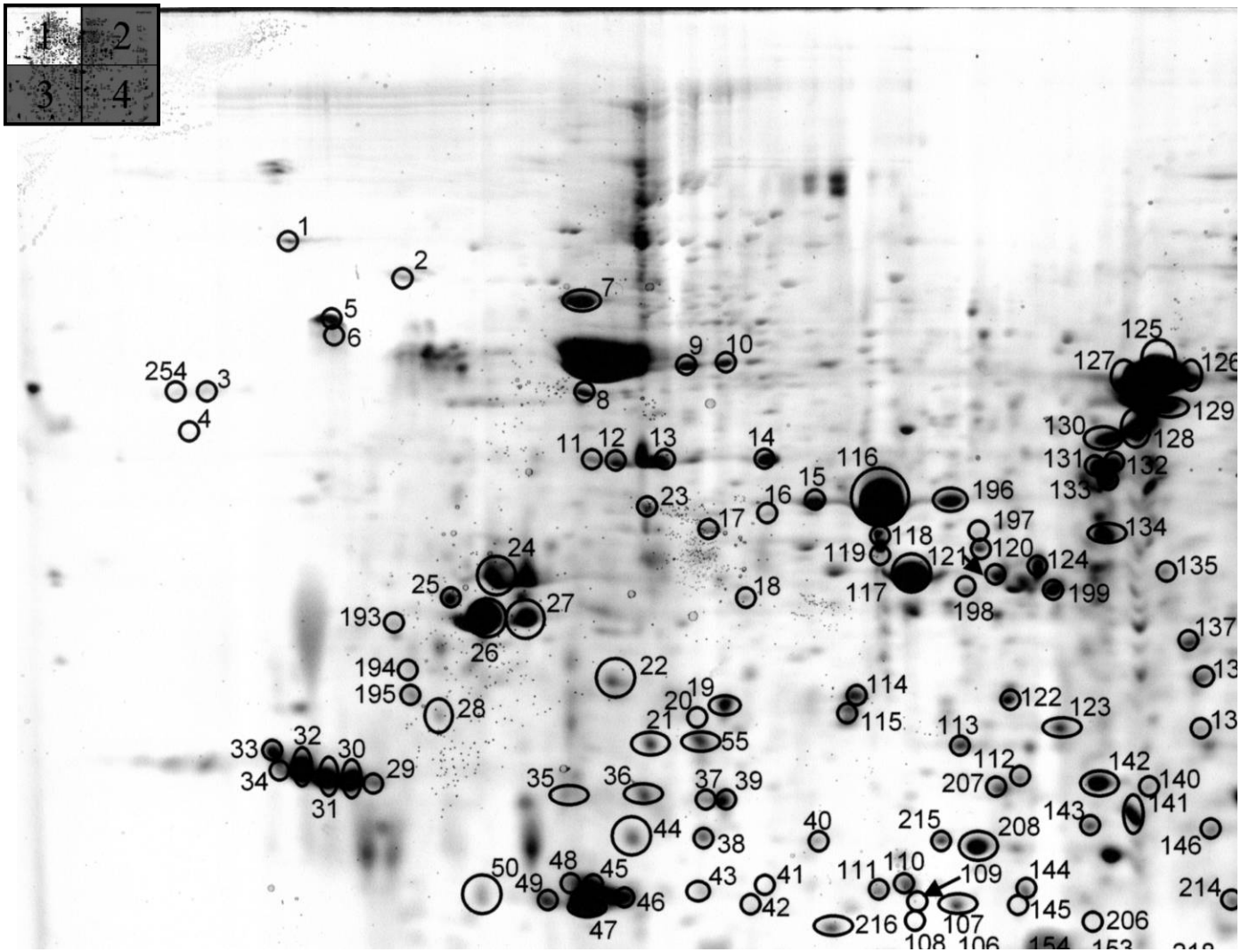


Fig. 14. Worker caste mass spectrometry reference map – Quadrant 1.
Inset: Highlight of selected gel quadrant.

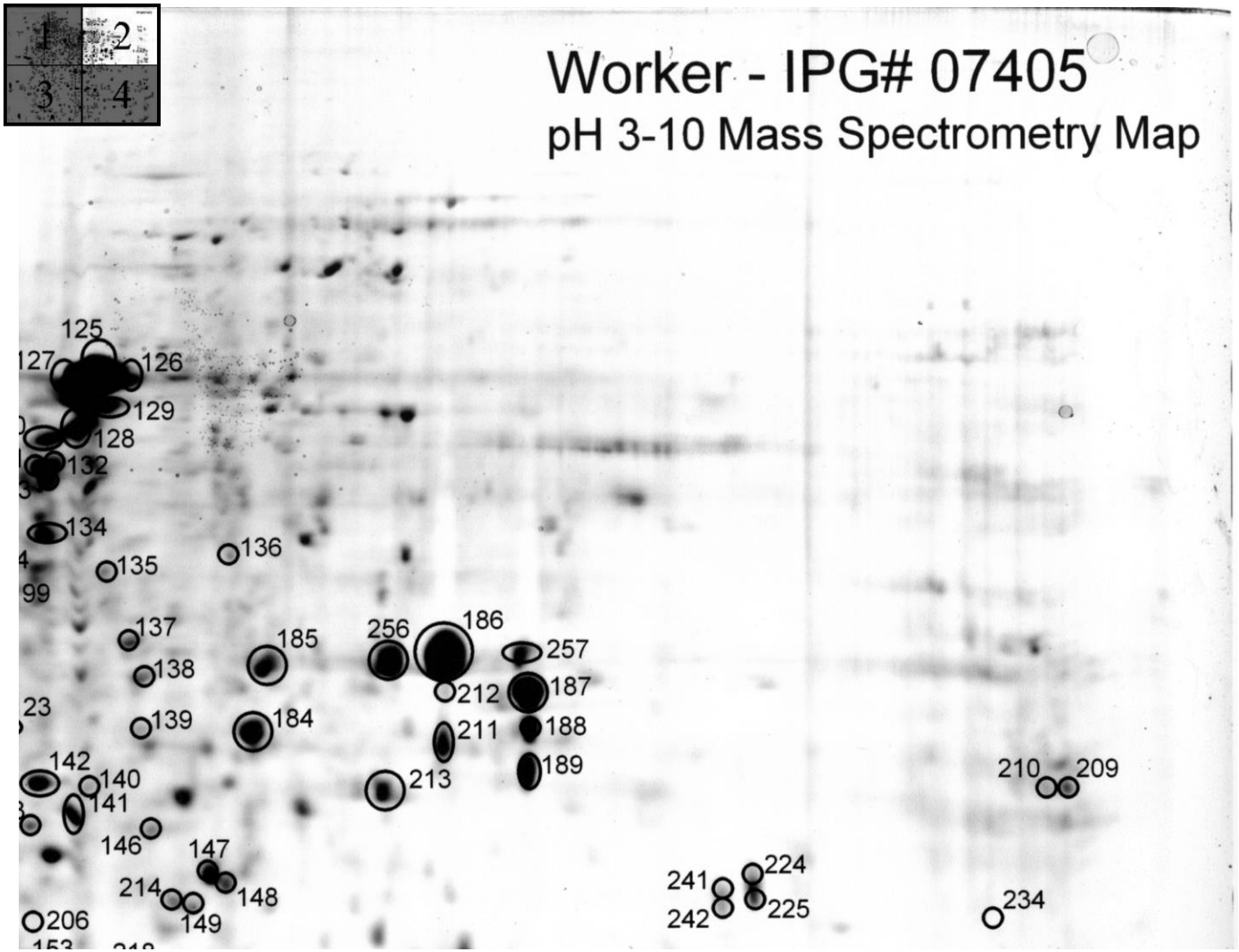


Fig. 15. Worker caste mass spectrometry reference map – Quadrant 2.
Inset: Highlight of selected gel quadrant.

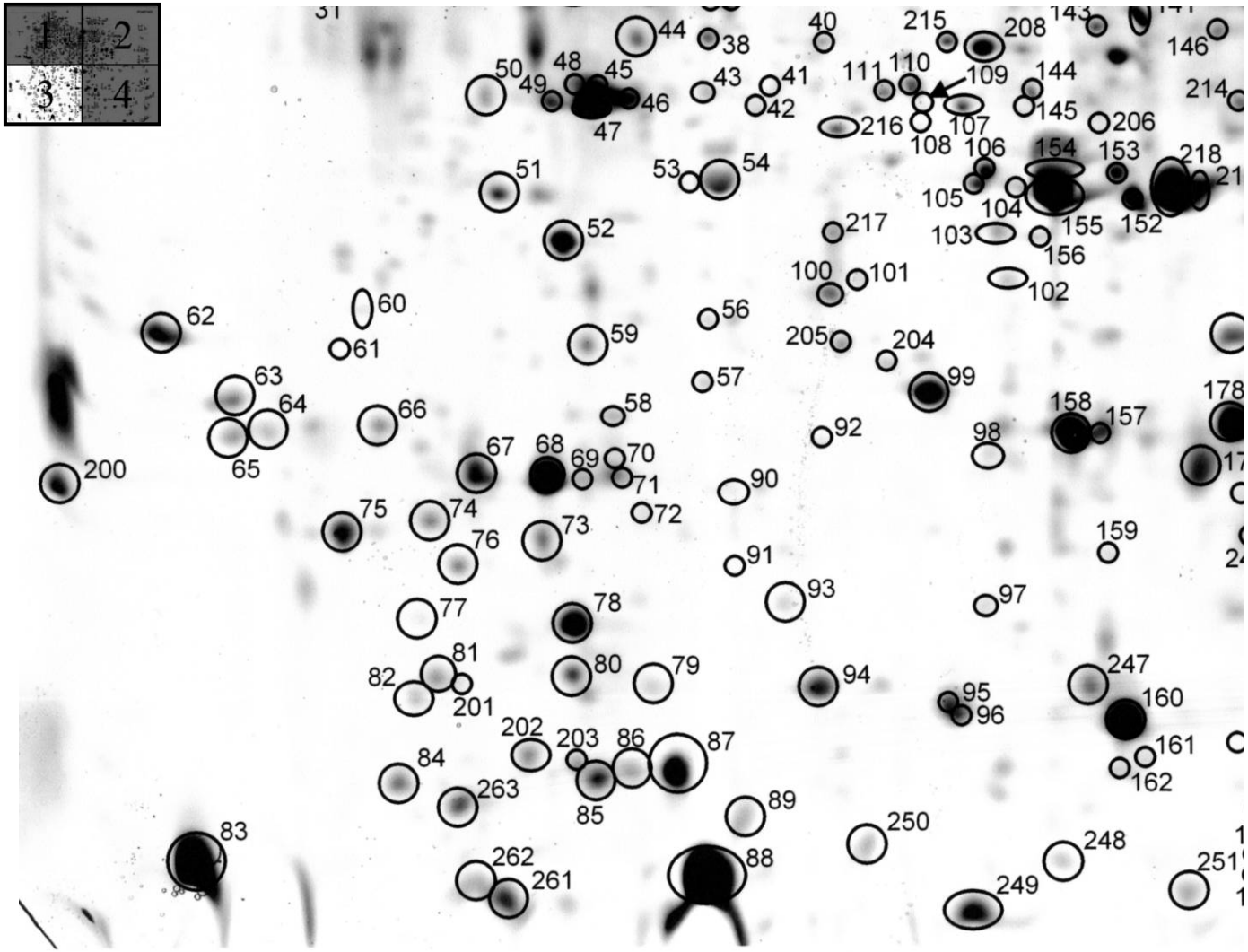


Fig. 16. Worker caste mass spectrometry reference map – Quadrant 3.
Inset: Highlight of selected gel quadrant.

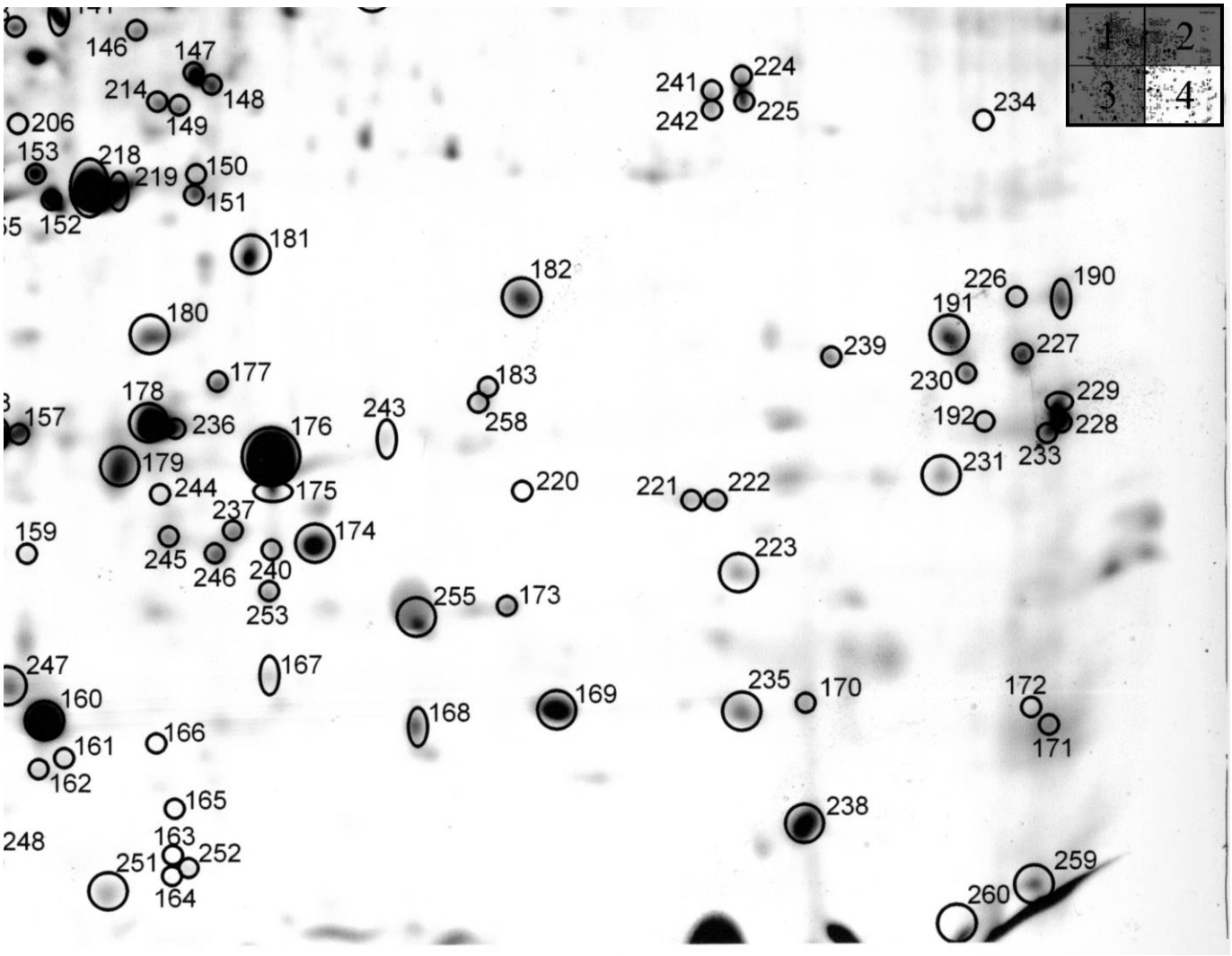


Fig. 17. Worker caste mass spectrometry reference map – Quadrant 4.
Inset: Highlight of selected gel quadrant.

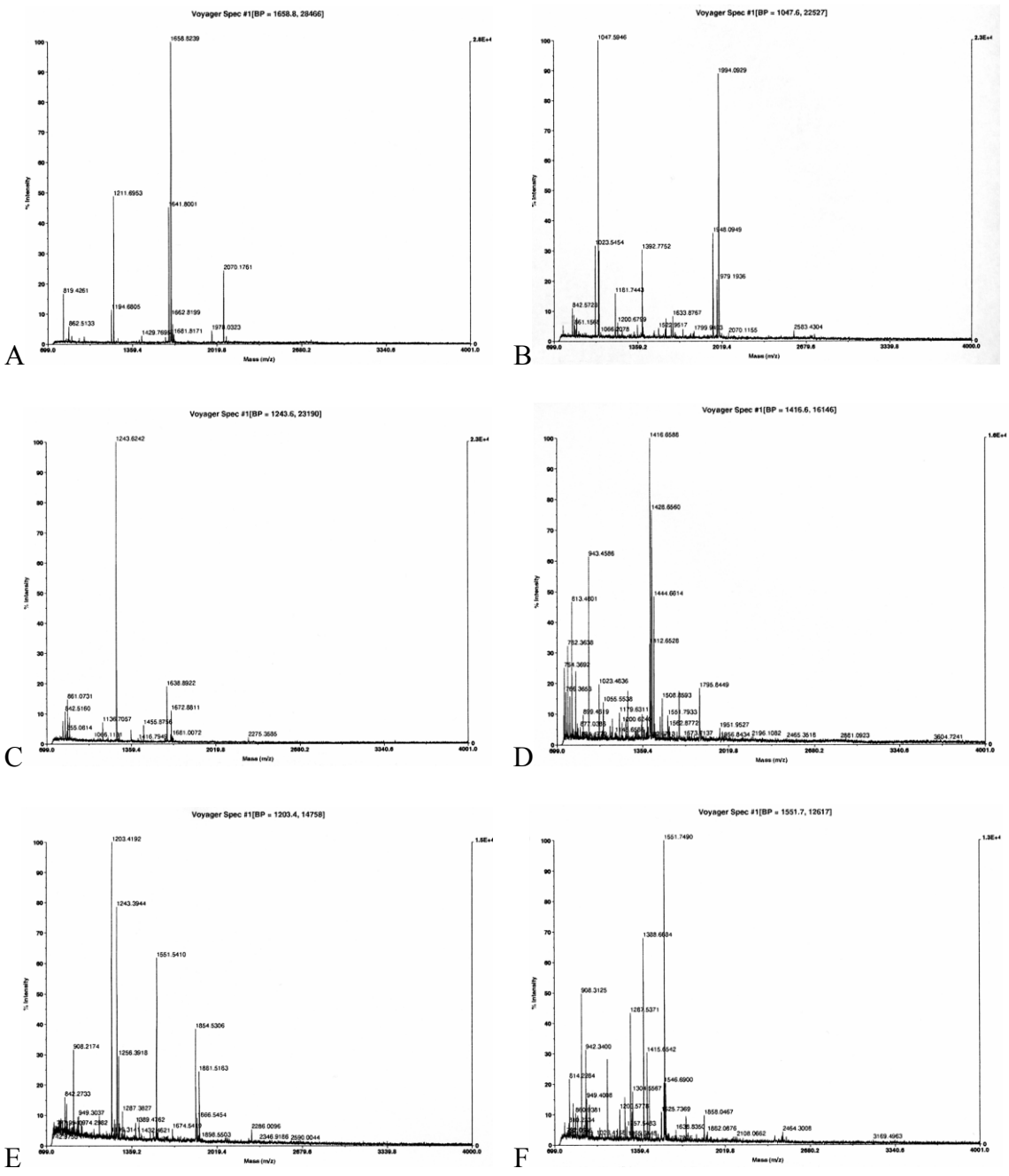
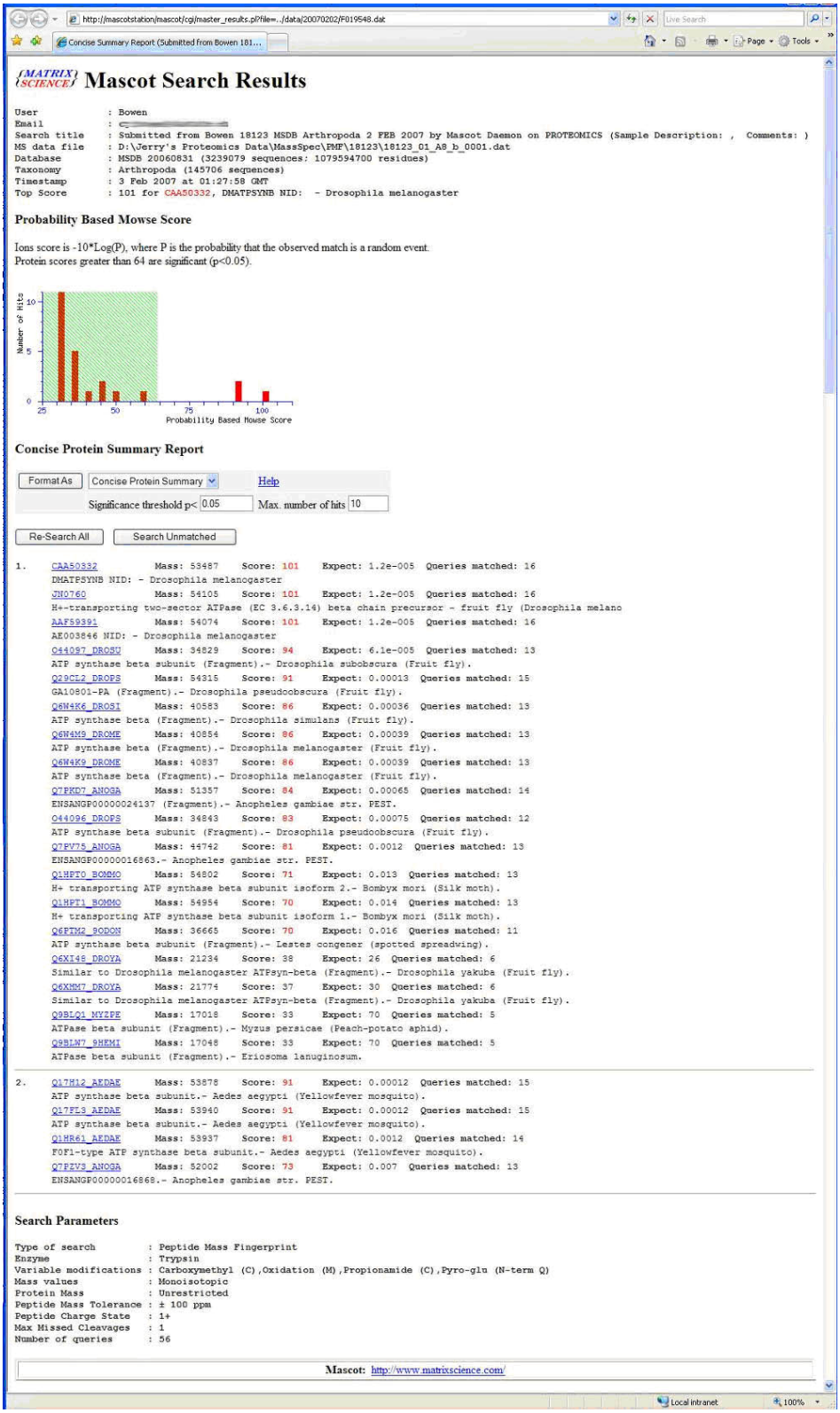


Fig. 18. Examples of MALDI-TOF generated peptide mass fingerprints: (A) putative arginine kinase; (B) coagulation factor B; (C) AF106961 NID; (D) DMATPSYNB NID; (E) putative muscle actin; (F) actin (clone 12).

(Refer to enclosed CD or web site (http://www.ento.okstate.edu/labs/jwd/index.htm) for complete listing of data including: spectra, peaklist, pI, MW, and related database search information.)



http://mascotstation/mascot/cgi/master_results.pl?file=../data/20070202/F019548.dat

Concise Summary Report (Submitted from Bowen 181...

Mascot Search Results

User : Bowen
Email : [redacted]
Search title : Submitted from Bowen 18123 MSDB Arthropoda 2 FEB 2007 by Mascot Daemon on PROTEOMICS (Sample Description: , Comments:)
MS data file : D:\Jerry's Proteomics Data\MassSpec\FMF\18123\18123_01_A8_b_0001.dat
Database : MSDB 20060831 (3239079 sequences; 1079594700 residues)
Taxonomy : Arthropoda (145706 sequences)
Timestamp : 3 Feb 2007 at 01:27:58 GMT
Top Score : 101 for CAA50332, DMATPSYNB NID: - Drosophila melanogaster

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 64 are significant (p<0.05).

Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 Max. number of hits 10
Re-Search All | Search Unmatched

1. CAA50332 Mass: 53487 Score: 101 Expect: 1.2e-005 Queries matched: 16
DMATPSYNB NID: - Drosophila melanogaster
JN0760 Mass: 54105 Score: 101 Expect: 1.2e-005 Queries matched: 16
H+-transporting two-sector ATPase (EC 3.6.3.14) beta chain precursor - fruit fly (Drosophila melano
AAF59391 Mass: 54074 Score: 101 Expect: 1.2e-005 Queries matched: 16
AE003846 NID: - Drosophila melanogaster
O44097_DROSU Mass: 34829 Score: 94 Expect: 6.1e-005 Queries matched: 13
ATP synthase beta subunit (Fragment).- Drosophila subobscura (Fruit fly).
Q29CL2_DROPS Mass: 54315 Score: 91 Expect: 0.00013 Queries matched: 15
GA10801-PA (Fragment).- Drosophila pseudoobscura (Fruit fly).
Q6W4K6_DROSI Mass: 40583 Score: 86 Expect: 0.00036 Queries matched: 13
ATP synthase beta (Fragment).- Drosophila simulans (Fruit fly).
Q6W4M9_DROME Mass: 40854 Score: 86 Expect: 0.00039 Queries matched: 13
ATP synthase beta (Fragment).- Drosophila melanogaster (Fruit fly).
Q6W4K9_DROME Mass: 40837 Score: 86 Expect: 0.00039 Queries matched: 13
ATP synthase beta (Fragment).- Drosophila melanogaster (Fruit fly).
Q7PKD7_ANOGA Mass: 51357 Score: 84 Expect: 0.00065 Queries matched: 14
ENSANGP00000024137 (Fragment).- Anopheles gambiae str. PEST.
O44096_DROPS Mass: 34843 Score: 83 Expect: 0.00075 Queries matched: 12
ATP synthase beta subunit (Fragment).- Drosophila pseudoobscura (Fruit fly).
Q7PV75_ANOGA Mass: 44742 Score: 81 Expect: 0.0012 Queries matched: 13
ENSANGP00000016863.- Anopheles gambiae str. PEST.
Q1HPT0_BOMMO Mass: 54802 Score: 71 Expect: 0.013 Queries matched: 13
H+ transporting ATP synthase beta subunit isoform 2.- Bombyx mori (Silk moth).
Q1HPT1_BOMMO Mass: 54954 Score: 70 Expect: 0.014 Queries matched: 13
H+ transporting ATP synthase beta subunit isoform 1.- Bombyx mori (Silk moth).
Q6PTM2_9ODON Mass: 36665 Score: 70 Expect: 0.016 Queries matched: 11
ATP synthase beta subunit (Fragment).- Lestes congener (spotted spreadwing).
Q6XI48_DROYA Mass: 21234 Score: 38 Expect: 26 Queries matched: 6
Similar to Drosophila melanogaster ATPsyn-beta (Fragment).- Drosophila yakuba (Fruit fly).
Q6XHM7_DROYA Mass: 21774 Score: 37 Expect: 30 Queries matched: 6
Similar to Drosophila melanogaster ATPsyn-beta (Fragment).- Drosophila yakuba (Fruit fly).
Q9BLQ1_MYZPE Mass: 17018 Score: 33 Expect: 70 Queries matched: 5
ATPase beta subunit (Fragment).- Myzus persicae (Peach-potato aphid).
Q9BLW7_9HEMI Mass: 17048 Score: 33 Expect: 70 Queries matched: 5
ATPase beta subunit (Fragment).- Eriosoma lanuginosum.

2. Q17H12_AEDAE Mass: 53878 Score: 91 Expect: 0.00012 Queries matched: 15
ATP synthase beta subunit.- Aedes aegypti (Yellowfever mosquito).
Q17FL3_AEDAE Mass: 53940 Score: 91 Expect: 0.00012 Queries matched: 15
ATP synthase beta subunit.- Aedes aegypti (Yellowfever mosquito).
Q1HR61_AEDAE Mass: 53937 Score: 81 Expect: 0.0012 Queries matched: 14
F0F1-type ATP synthase beta subunit.- Aedes aegypti (Yellowfever mosquito).
Q7PZV3_ANOGA Mass: 52002 Score: 73 Expect: 0.007 Queries matched: 13
ENSANGP00000016868.- Anopheles gambiae str. PEST.

Search Parameters

Type of search : Peptide Mass Fingerprint
Enzyme : Trypsin
Variable modifications : Carboxymethyl (C),Oxidation (M),Propionamide (C),Pyro-glu (N-term Q)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 100 ppm
Peptide Charge State : 1+
Max Missed Cleavages : 1
Number of queries : 56

Mascot: http://www.matrixscience.com/

Local intranet 100%

Fig. 19. Example of database searching – search results. (www.matrixscience.com)

(Refer to enclosed CD or web site (http://www.ento.okstate.edu/labs/jwd/index.htm) for complete listing of data including: spectra, peaklist, pI, MW, and related database search information.)

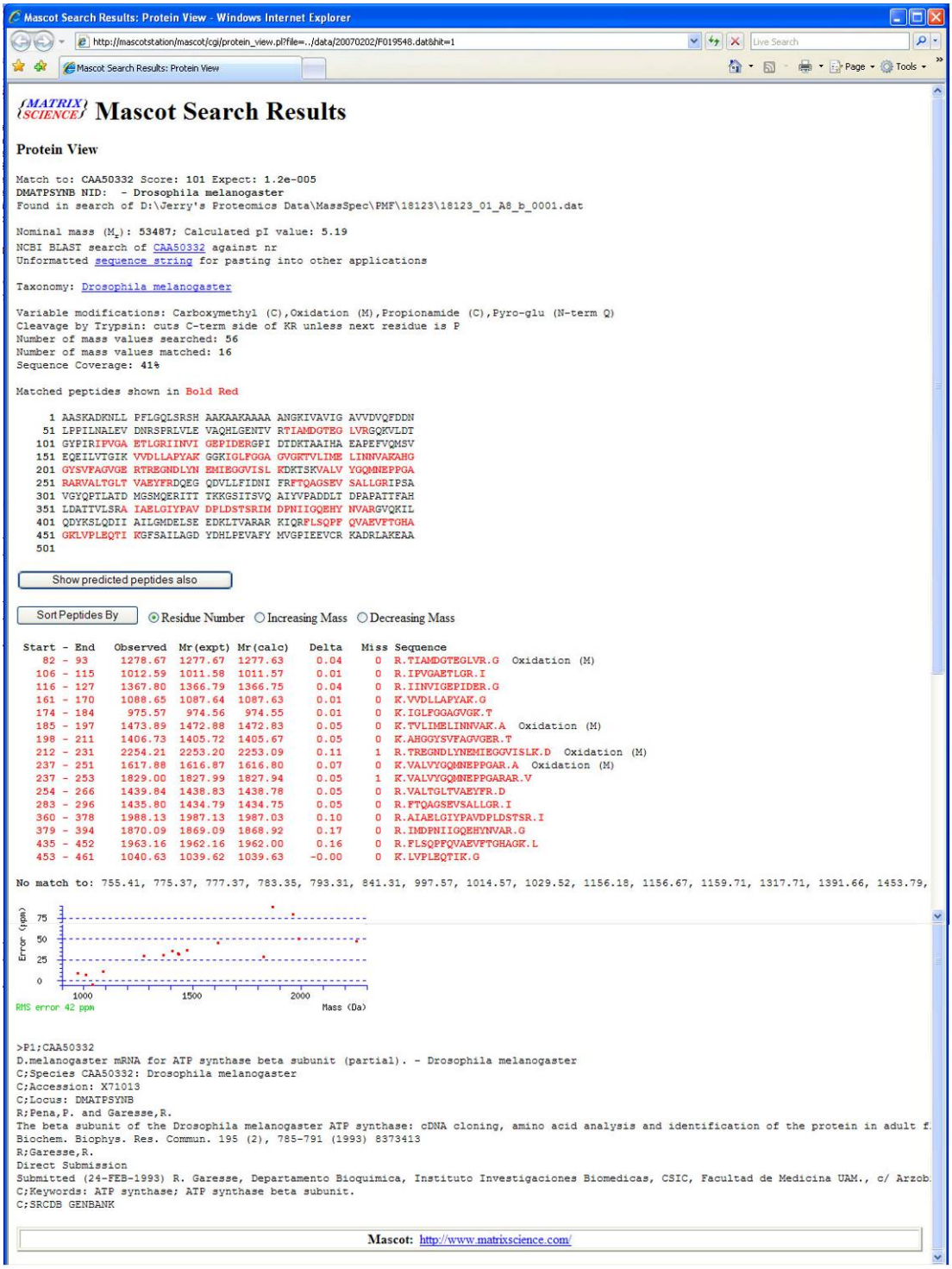


Fig. 20. Example of Database Searching – protein view. (www.matrixscience.com)

(Refer to enclosed CD or web site (http://www.ento.okstate.edu/labs/jwd/index.htm) for complete listing of data including: spectra, peaklist, pI, MW, and related database search information.)

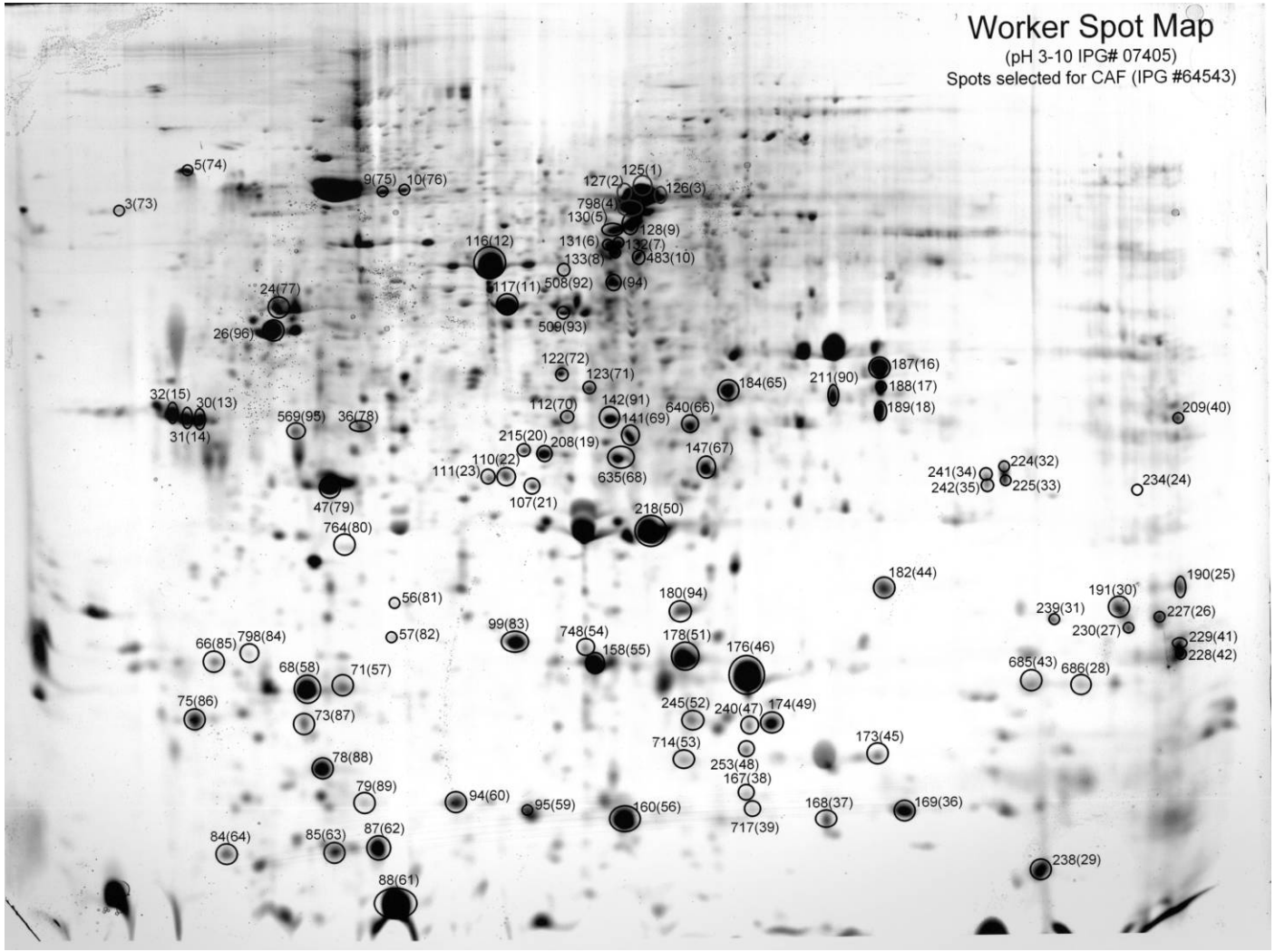


Fig. 21. Worker caste CAF mass spectrometry reference map for *Reticulitermes flavipes*.

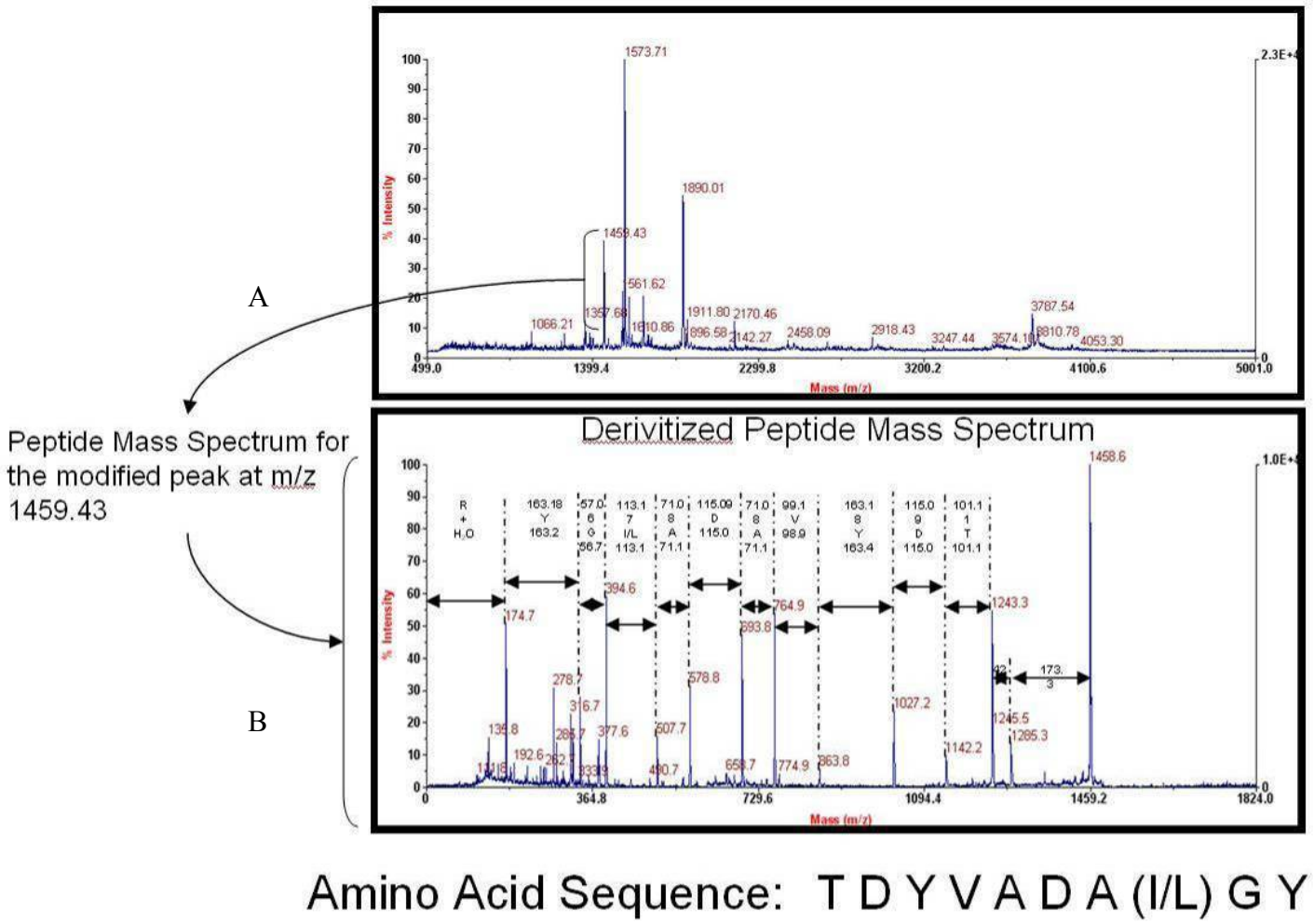


Fig. 22. Example of spectra for CAF: (A) Peptide mass fingerprint and (B) peptide mass fingerprint for the modified peak at *m/z* 1459.43. Amino acid sequence was elucidated as TDYVADA(I/L)GY. Putative identification was 'cuticular protein'.

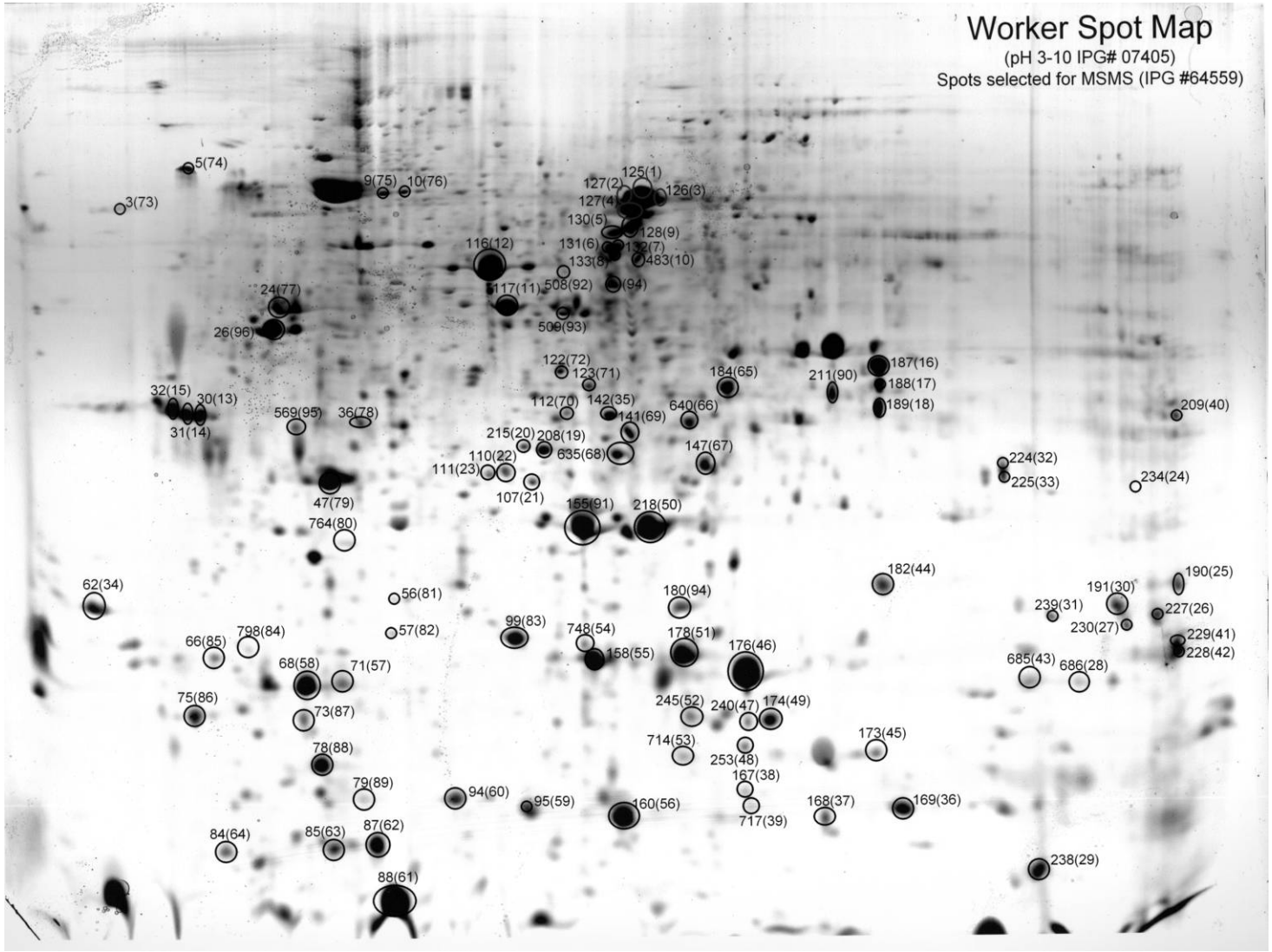
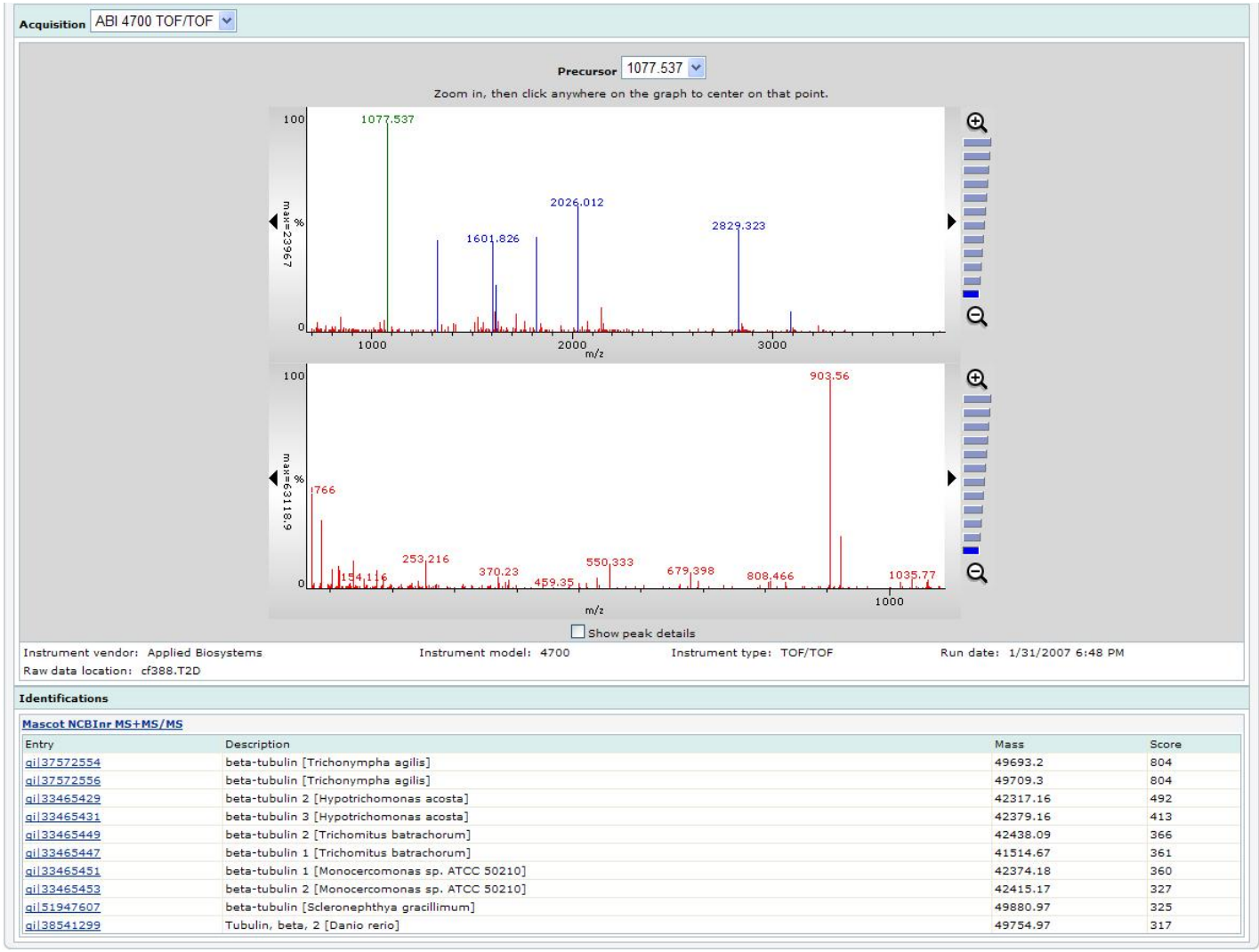


Fig. 23. Worker caste MS/MS mass spectrometry reference map for *Reticulitermes flavipes*.



Entry	Description	Mass	Score
gi|37572554	beta-tubulin [Trichonympha agilis]	49693.2	804
gi|37572556	beta-tubulin [Trichonympha agilis]	49709.3	804
gi|33465429	beta-tubulin 2 [Hypotrichomonas acosta]	42317.16	492
gi|33465431	beta-tubulin 3 [Hypotrichomonas acosta]	42379.16	413
gi|33465449	beta-tubulin 2 [Trichomitus batrachorum]	42438.09	366
gi|33465447	beta-tubulin 1 [Trichomitus batrachorum]	41514.67	361
gi|33465451	beta-tubulin 1 [Monocercomonas sp. ATCC 50210]	42374.18	360
gi|33465453	beta-tubulin 2 [Monocercomonas sp. ATCC 50210]	42415.17	327
gi|51947607	beta-tubulin [Scleronephthya gracillimum]	49880.97	325
gi|38541299	Tubulin, beta, 2 [Danio rerio]	49754.97	317

Fig. 24. Example of peptide mass spectra and results from MS/MS database search.

Table 4. Putative identifications with a 'termite' source

Confidence	Map Spot #	MS Type	Exp. pI	Exp. Mass	Protein Identifier	Score	Expect	Description	Source
high	32	MS/MS	3.87	29884	RFW0387_029884	195	9.8e-014	hypothetical protein	*Hodotermopsis sjoestedti*
high	62	MS/MS	3.41	17625	RFW0341_017625	112	2e-005	hexamerin II	*Reticulitermes flavipes*
high	87	MS/MS	5.00	7326	RFW0500_007326	90	0.0028	hexamerin I	*Reticulitermes flavipes*
high	95	MS/MS	5.92	8340	RFW0592_008340	140	3.1e-008	hexamerin II	*Reticulitermes flavipes*
high	117	MS/MS	5.16	45647	RFW0516_045647	216	7.8e-016	hypothetical protein	*Hodotermopsis* sp.
high	130	MS/MS	5.45	66430	RFW0545_066430	377	6.2e-032	hexamerin II	*Reticulitermes flavipes*
high	142	MS/MS	5.92	30992	RFW0592_030992	85	0.0094	hexamerin I	*Reticulitermes flavipes*
high	158	MS/MS	6.30	18089	RFW0630_018089	58	4.6	hexamerin II	*Reticulitermes flavipes*
high	167	MS/MS	7.14	9069	RFW0714_009069	195	9.8e-014	hexamerin I	*Reticulitermes flavipes*
high	169	MS/MS	8.04	8387	RFW0767_007550	131	2.5e-007	hexamerin I	*Reticulitermes flavipes*
high	225	MS/MS	8.65	27104	RFW0865_027104	63	1.6	hexamerin II	*Reticulitermes flavipes*
high	238	MS/MS	8.83	6724	RFW0883_006724	174	1.2e-011	hexamerin I	*Reticulitermes flavipes*
high	635	MS/MS	6.03	29396	RFW0603_029396	82	0.019	hexamerin II	*Reticulitermes flavipes*
good	73	PMF	4.61	14460	RFW0461_014460	48	2.8	hypothetical protein	*Nasutitermes takasogoensis*
good	253	MS/MS	7.14	10551	RFW0714_010551	36	7.1e+002	hexamerin I	*Reticulitermes flavipes*
low	648	PMF	6.18	27299	RFW0618_027299	48	3	family 4 cytochrome P450	*Coptotermes acinaciformis*

End of Table

Table 5. Putative worker protein identifications

(Refer to enclosed CD or web site (http://www.ento.okstate.edu/labs/jwd/index.htm) for complete listing of data including: spectra, peaklist, pI, MW, and related database search information.)

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	1	PMF	3.74	89065	RFW0374_089065	43	8	major royal jelly protein 2	*Apis mellifera*
low	2	PMF	3.94	89642	RFW0394_089642	47	3.3	ENSANGP00000018491	*Anopheles gambiae* str. PEST
low	3	MS/MS	3.77	72688	RFW0377_072688	74	0.13	PREDICTED: similar to ENSANGP00000020389	*Apis mellifera*
low	4	PMF	3.49	67250	RFW0349_067250	38	25	coagulation factor B precursor	*Tachypleus tridentatus*
low	5	PMF	4.01	67844	RFW0401_067844	37	39	GA11336-PA	*Drosophila pseudoobscura*
high	7	PMF	4.54	97308	RFW0454_097308	78	0.0027	Myosin heavy chain CG17927-PM, isoform M	*Drosophila melanogaster*
high	8	PMF	4.44	81256	RFW0444_081256	101	1.4e-005	ATP synthase beta subunit	*Drosophila melanogaster*
high	9	MS/MS	4.72	79321	RFW0472_079321	77	0.064	glutathione-S-transferase-like protein	*Galleria mellonella*
high	10	MS/MS	4.86	77426	RFW0486_077426	88	0.0048	glutathione-S-transferase-like protein	*Galleria mellonella*
low	11	PMF	4.59	53883	RFW0459_053883	49	2.4	PREDICTED: similar to Nuclear pore complex protein Nup160 homolog	*Apis mellifera*
high	12	PMF	4.63	54335	RFW0463_054335	156	4.4E-011	Actin, clone 211	*Artemia* sp.
high	13	PMF	4.61	56794	RFW0461_056794	85	0.00057	putative muscle actin	*Homalodisca coagulata* (Glassy-winged sharpshooter)
moderate	14	PMF	4.81	64668	RFW0481_064668	54	0.64	PREDICTED: similar to Inositol 1,4,5,-tris-phosphate receptor CG1063-PB, isoform B	*Apis mellifera*
high	15	PMF	4.95	56644	RFW0495_056644	79	0.0023	creatine kinase	*Aedes aegypti* (Yellowfever mosquito)
moderate	16	PMF	5.35	57285	RFW0535_057285	52	1	Actin, clone 211	*Artemia* sp.
high	17	PMF	4.87	47682	RFW0487_047682	82	0.0012	beta-2 tubulin	*Laodelphax striatellus*
high	18	PMF	5.19	41602	RFW0519_041602	98	3e-005	putative actin	*Diaphorina citri*
high	19	PMF	5.30	34612	RFW0530_034612	97	3.7e-005	Putative arginine kinase	*Homolodisca coagulata*
good	20	PMF	5.20	34182	RFW0520_064182	61	0.14	ENSANGP00000012828	*Anopheles gambiae* str. PEST
moderate	21	PMF	5.11	31950	RFW0511_031950	59	0.22	CAST	*Drosophila melanogaster*
low	22	PMF	4.99	35120	RFW0499_035120	49	2.4	PREDICTED: similar to CG16916-PA	*Tribolium castaneum*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	23	PMF	4.75	49666	RFW0475_049666	50	1.9	4-nitrophenylphosphatase	*Aedes aegypti* (Yellowfever mosquito)
high	24	PMF	4.49	51451	RFW0449_051451	112	1.1e-006	Tropomyosin	*Periplaneta americana*
moderate	25	PMF	4.48	39318	RFW0448_039318	54	0.71	GA13813-PA	*Drosophila pseudoobscura*
low	26	PMF	4.56	38734	RFW0456_038734	50	1.9	PREDICTED: similart to CF3886-PA	*Tribolium castaneum*
low	27	PMF	4.69	39305	RFW0469_039305	45	5.8	ENSANGP00000017422	*Anopheles gambiae* str. PEST
low	28	PMF	4.46	33296	RFW0446_033296	44	7.4	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	29	PMF	4.12	34327	RFW0412_034327	44	7.7	PREDICTED: similar to CG30337-PB, isoform B	*Tribolium castaneum*
high	30	MS/MS	4.05	29393	RFW0405_029393	804	1.2e-074	beta-tubulin	*Trichonympha agilis*
high	31	MS/MS	3.97	29682	RFW0397_039682	81	0.027	peroxiredoxin-like protein	*Aedes aegypti*
high	32	MS/MS	3.87	29884	RFW0387_029884	195	9.8e-014	hypothetical protein	*Hodotermopsis sjoestedti*
low	33	PMF	3.77	29998	RFW0377_029998	45	5.7	CG9018-PB, isoform B	*Drosophila melanogaster*
low	34	PMF	3.78	35373	RFW0378_035373	43	8.9	CG14434-PA	*Drosophila melanogaster*
low	35	PMF	4.91	28512	RFW0491_028512	38	29	PIWI	*Aedes aegypti* (Yellowfever mosquito)
high	36	PMF	5.12	29071	RFW0512_029071	80	0.0017	arginine kinase	*Periplaneta americana*
high	37	PMF	5.14	33191	RFW0514_033191	65	0.05	DEAD box ATP-dependent RNA helicase	*Aedes aegypti* (Yellowfever mosquito)
low	38	PMF	5.28	27879	RFW0528_027879	47	3.6	cytochrome P450	*Drosophila simulans*
low	39	PMF	5.35	29521	RFW0535_029521	30	1.9e+002	adenylsulfate kinase	*Aedes aegypti* (Yellowfever mosquito)
low	40	PMF	5.62	26802	RFW0562_026802	33	90	PREDICTED: similar to CG17052-PA	*Apis mellifera*
moderate	41	PMF	5.48	25449	RFW0548_025449	55	0.59	hypothetical protein Aael_AAEL002036	*Aedes aegypti* (Yellowfever mosquito)
low	42	PMF	5.43	24682	RFW0543_024682	36	49	Cardioactive peptide precursor (CCAP) (Crusatcean cardioactive peptide)	*Periplaneta americana*
low	43	PMF	5.31	24878	RFW0531_024878	48	2.6	putative transposase yabusame-1	*Bombyx mori*
good	44	PMF	5.08	26522	RFW0508_026522	61	0.15	DEAD-box protein abstrakt	*Drosophila melanogaster*
moderate	45	PMF	5.02	24972	RFW0502_024972	55	0.59	multiprotein bridging factor, putative	*Aedes aegypti* (Yellowfever mosquito)
moderate	46	PMF	4.89	27318	RFW0489_027318	53	0.85	Porin2 CG17137-PA	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
high	47	PMF	4.94	24245	RFW0494_024245	83	0.00087	heat shock protein 20.6	*Locusta migratoria*
good	48	PMF	4.73	28240	RFW0473_028240	56	0.46	PREDICTED: similar to Recombination repair protein 1 (DNA- (apurinic or apyrimidinic site) lyase)	*Apis mellifera*
low	49	PMF	4.65	27120	RFW0465_027120	39	21	GM05777p	*Drosophila melanogaster*
low	50	PMF	4.69	24164	RFW0469_024164	47	3.7	ENSANGP00000015666	*Anopheles gambiae* str. PEST
good	51	PMF	4.47	21387	RFW0447_021387	57	0.37	GA17230-PA	*Drosophila pseudoobscura*
low	52	PMF	4.68	20266	RFW0468_020266	52	1.1	IP09809p	*Drosophila melanogaster*
low	53	PMF	5.04	22913	RFW0504_022913	51	1.5	RE20606p	*Drosophila melanogaster*
low	54	PMF	5.37	22535	RFW0537_022535	46	4.2	reverse transcriptase	*Drosophila ambigua*
good	56	PMF	5.14	17834	RFW0514_017834	60	0.17	SD23787p	*Drosophila melanogaster*
moderate	57	PMF	5.43	17673	RFW0543_017676	52	1	GA17940-PA	*Drosophila pseudoobscura*
moderate	58	PMF	4.87	16732	RFW0487_016732	54	0.66	ribosomal protein 14	*Lonomia obliqua*
low	59	PMF	4.74	18141	RFW0474_018141	47	3.7	Liprin-alpha CG11199-PA, isoform A	*Drosophila melanogaster*
low	60	PMF	4.03	18234	RFW0403_018234				
moderate	61	PMF	4.00	16674	RFW0400_016674	63	0.097	similar to Drosophila melanogaster betaTub56D	*Drosophila yakuba* (Fruit fly)
high	62	MS/MS	3.41	17625	RFW0341_017625	112	2e-005	hexamerin II	*Reticulitermes flavipes*
low	63		3.63	16285	RFW0363_016285				
low	64	PMF	3.75	14409	RFW0375_014409	27	3.4e+002	PREDICTED: similar to Ras-like protein 1	*Apis mellifera*
low	65	PMF	3.64	14302	RFW0364_014302	34	75	ribosomal protein S27e	*Agriotes lineatus*
low	66	MS/MS	4.10	16080	RFW0410_016080	41	2.5e+002	arginine kinase	*Epicephala* sp. E38AT
low	67	PMF	4.41	15485	RFW0441_015485	31	1.3e+002	hypothetical protein Aael_AAEL005474	*Aedes aegypti* (Yellowfever mosquito)
low	68	PMF	4.67	15450	RFW0467_015450	43	8.2	CG1129-PA, isoform A	*Drosophila melanogaster*
low	69	PMF	4.75	13008	RFW0475_013008	37	35	orfY, putative	*Aedes aegypti* (Yellowfever mosquito)
low	70	PMF	4.84	13764	RFW0484_013764	31	1.5e+002	LD25702p	*Drosophila melanogaster*
low	71	PMF	4.88	15811	RFW0488_015811	44	6.9	TPA: TPA_inf: HDC17852	*Drosophila melanogaster*
low	72	PMF	4.95	14976	RFW0495_014976	36	45	CG1969-PB, isoform B	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
high	73	MS/MS	4.61	14460	RFW0461_014460	119	3.9e-006	beta actin	*Hippoglossu hippoglossus*
low	74	PMF	4.25	14593	RFW0425_014593	29	2.3e+002	TPA: TPA_inf: HDC12287	*Drosophila melanogaster*
low	75	PMF	3.98	14401	RFW0398_014401	49	2.4	Tubby, putative	*Aedes aegypti* (Yellowfever mosquito)
moderate	76	PMF	4.37	14137	RFW0437_014137	53	0.97	PREDICTED: similar to Malate dehydrogenase, mitochondrial precursor	*Tribolium castaneum*
low	77		4.22	13412	RFW0422_013412				
low	78	MS/MS	4.70	13319	RFW0470_013319	51	25	actin	*Culex pipeins pipiens*
low	79	PMF	4.99	12602	RFW0499_012602	42	12	ENSANGP00000020183	*Anopheles gambiae* str. PEST
low	80	PMF	4.72	12704	RFW0472_012704	46	4.1	PREDICTED: similar to CG10225-PA	*Tribolium castaneum*
low	81	PMF	4.29	12749	RFW0429_012749	29	2.2e+002	PREDICTED: hypothetical protein	*Apis mellifera*
low	82	PMF	4.20	12431	RFW0420_012431	27	3.4e+002	toll-related protein	*Aedes aegypti* (Yellowfever mosquito)
low	83	PMF	3.58	10573	RFW0358_010573	33	84	CG32147-PA	*Drosophila melanogaster*
high	84	MS/MS	4.15	11457	RFW0415_011457	119	3.9e-006	ATP synthase alpha chain-like protein	*Magnaporthe grisea*
low	85	PMF	4.79	7216	RFW0479_007216	43	9	PREDICTED: similar to CG8597-PA, isoform A	*Tribolium castaneum*
low	86	PMF	4.88	7467	RFW0488_007467	34	63	GA18132-PA	*Drosophila pseudoobscura*
high	87	MS/MS	5.00	7326	RFW0500_007326	90	0.0028	hexamerin I	*Reticulitermes flavipes*
low	88	MS/MS	5.16	6220	RFW0516_006220	37	5.7e+002	PREDICTED: similar to hypothetical protein FLJ11767	*Gallus gallus*
low	89	PMF	5.24	6889	RFW0524_006889	27	3.2e+002	Nitrophorin-2 precursor (NP2) (Prolixin-S)	*Rhodnius prolixus*
low	90	PMF	5.52	15871	RFW0552_015871	48	2.8	giant secretory protein	*Chironomus pallidivittatus*
low	91	PMF	5.57	14940	RFW0557_014940	48	2.8	PREDICTED: similar to CG4058-PA, isoform A	*Tribolium castaneum*
low	92	PMF	5.73	16747	RFW0573_016747	32	1.1e+002	PREDICTED: similar to CG6294-PA	*Tribolium castaneum*
low	93	PMF	5.71	13952	RFW0571_013952	39	21	ENSANGP00000019534	*Anopheles gambiae* str. PEST
high	94	MS/MS	5.85	12752	RFW0585_012752	104	0.00012	ubiquitin	*Homo sapiens*
high	95	MS/MS	5.92	8340	RFW0592_008340	140	3.1e-008	hexamerin II	*Reticulitermes flavipes*
low	96	PMF	5.92	8309	RFW0592_008309	28	2.7e+002	putative accessory gland protein	*Gryllus veletis*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	97	PMF	6.29	13826	RFW0629_013826	38	31	CG13900-PB, isoform B	*Drosophila melanogaster*
low	98	PMF	6.03	17694	RFW0603_017694	51	1.3	GA13054-PA	*Drosophila pseudoobscura*
low	99	PMF	5.99	18483	RFW0599_018483	46	4.7	ENSANGP00000016550	*Anopheles gambiae* str. PEST
moderate	100	PMF	5.73	20008	RFW0573_020008	53	0.93	spellchecker1 CG4215-PC, isoform C	*Drosophila melanogaster*
low	101	PMF	5.82	20524	RFW0582_020524	25	5.9e+002	PREDICTED: similar to CG5731-PA	*Tribolium castaneum*
low	102	PMF	6.08	21735	RFW0608_021735	48	2.6	CG1490	*Drosophila miranda*
low	103	PMF	6.08	22813	RFW0608_022813	45	5	ENSANGP00000016796	*Anopheles gambiae* str. PEST
low	104	PMF	6.13	22942	RFW0613_022942	51	1.3	PREDICTED: similar to DnaJ (Hsp40) homolog, subfamily C, member 13	*Tribolium castaneum*
high	105	PMF	5.95	23523	RFW0595_023523	67	0.038	origin recognition complex subunit	*Aedes aegypti* (Yellowfever mosquito)
low	106	PMF	6.07	23895	RFW0607_023895	43	9	mitochondrial NADH: ubiquinone oxidoreductase ESSS subunit, putative	*Aedes aegypti* (Yellowfever mosquito)
low	107	PMF	6.05	25156	RFW0605_025156	50	1.9	Rnase H and integrase-like protein	*Bombyx mori*
low	108	PMF	5.81	25934	RFW0581_025934	45	5.9	ENSANGP00000030095	*Anopheles gambiae* str. PEST
low	109	PMF	5.81	26958	RFW0581_026958	19	2e+003	IP15747p	*Drosophila melanogaster*
good	110	PMF	5.91	25555	RFW0591_025555	57	0.37	EcR	*Panorpa germanica*
good	111	PMF	5.81	25476	RFW0581_025476	59	0.24	Tropomyosin (Allergen Chi k 10)	*Chironmus kiiensis*
good	112	PMF	5.96	29632	RFW0596_029632	56	0.4	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	113	PMF	5.95	36640	RFW0595_036640	45	5.2	PREDICTED: hypothetical protein isoform 2	*Apis mellifera*
low	114	PMF	5.54	35649	RFW0554_035649	42	12	cytochrome P450 CYP4	*Cherax quadricarinatus*
good	115	PMF	5.61	39288	RFW0561_039288	60	0.17	PREDICTED: similar to CG11199-PA, isoform A	*Tribolium castaneum*
high	116	PMF	5.17	54053	RFW0517_054053	76	0.0044	arginine kinase	*Periplaneta americana*
high	117	MS/MS	5.16	45647	RFW0516_045647	216	7.8e-016	hypothtical protein	*Hodotermopsis* sp.
high	118	PMF	5.11	49942	RFW0511_049942	81	0.0013	putative argine kinase	*Homalodisca coagulata* (Glassy-winged sharpshooter)
good	119	PMF	5.08	49483	RFW0508_049483	59	0.22	PREDICTED: similar to CG16916-PA	*Tribolium castaneum*
low	120	PMF	5.24	44960	RFW0524_044960	37	35	GA14865-PA	Drosophila *pseudoobscura*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	121	PMF	6.09	50838	RFW0609_050838	40	17	CG9895-PA	*Drosophila virilis*
high	122	PMF	5.71	33979	RFW0571_033979	115	5.5e-007	Putative arginine kinase	*Oncometopia nigricans*
low	123	PMF	5.86	31742	RFW0586_031742	50	1.6	Gustatory receptor 93a, putative	*Aedes aegypti* (Yellowfever mosquito)
low	124	PMF	5.38	41226	RFW0538_041226	34	69	cysteine synthase	*Aedes aegypti* (Yellowfever mosquito)
high	125	MS/MS	6.13	22942	RFW0613_022942	66	0.78	PREDICTED: similar to tropomyosin 1	*Apis mellifera*
low	126	PMF	5.83	74663	RFW0583_074663	51	1.4	26.7kDa salivary protein	*Phlebotomus duboscqi*
low	127	PMF	6.50	76258	RFW0650_076258	48	2.6	CG14618-PA	*Drosophila melanogaster*
low	128	PMF	6.51	67289	RFW0651_067289	47	3.8	PREDICTED: similar to RNA polymerase I subunit CG10122-PA	*Apis mellifera*
low	129	PMF	5.74	71500	RFW0574_071500	47	3.3	NADH dehydrogenase subunit 3	*Amblyomma triguttatum*
high	130	MS/MS	5.45	66430	RFW0545_066430	377	6.2e-032	hexamerin II	*Reticulitermes flavipes*
low	131	PMF	5.39	44843	RFW0539_044843	51	14	Cyp304a1 CG7241-PA	*Drosophila melanogaster*
high	132	PMF	6.56	63552	RFW0656_063552	87	0.00036	arginine kinase	*Periplaneta americana*
high	133	MS/MS	6.55	60419	RFW0655_060419	113	1.6e-005	muscular protein 20	*Cicindela nimuta*
low	134	PMF	5.41	40788	RFW0541_040788	47	3.6	cytochrome P450	*Drosophila simulans*
low	135	PMF	5.57	38534	RFW0557_038534	51	1.5	mitochondrial ribosomal protein S34 CG13037-PA	*Drosophila melanogaster*
low	136	PMF	6.01	53085	RFW0601_053085	64	0.066	PREDICTED: similar to Centromeric protein E (CENP-E protein)	*Apis mellifera*
moderate	137	PMF	5.75	36535	RFW0575_036535	55	0.59	hypothetical protein AaeL_AAEL002036	*Aedes aegypti* (Yellowfever mosquito)
low	138	PMF	5.85	35393	RFW0585_035393	50	1.7	Tpi	*Drosophila jambulina*
low	139	PMF	5.98	33947	RFW0598_033947	51	1.5	TPA: TPA_inf: HDC10365	*Drosophila melanogaster*
good	140	PMF	5.90	31868	RFW0590_031868	61	0.14	ENSANGP00000011251	*Anopheles gambiae* str. PEST
good	141	PMF	5.96	30693	RFW0596_030693	61	0.15	IP04554p	*Drosophila melanogaster*
high	142	MS/MS	5.92	30992	RFW0592_030992	85	0.0094	hexamerin I	*Reticulitermes flavipes*
low	143	PMF	6.36	31449	RFW0636_031449	43	8.7	Twister CG10210-PA	*Drosophila melanogaster*
low	144	PMF	6.12	26653	RFW0612_026653	40	17	short-chain dehydrogenase/reductase	*Apis mellifera*
low	145	PMF	6.13	26557	RFW0613_026557	38	27	PREDICTED: similar to glaikit CG8825-PA	*Apis mellifera*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
good	146	PMF	6.14	30908	RFW0614_030908	58	0.27	arginine kinase	*Periplaneta americana*
moderate	147	PMF	6.40	29766	RFW0640_029766	52	1	CG11156	*Drosophila simulans*
low	148	PMF	6.49	29195	RFW0649_029195	51	1.3	mitochondrial brown fat uncoupling protein	*Aedes aegypti* (Yellowfever mosquito)
good	149	PMF	6.40	28792	RFW0640_028792	60	0.19	cell division control protein	*Aedes aegypti* (Yellowfever mosquito)
low	150	PMF	6.93	24032	RFW0693_024032	49	2	GA17547-PA	*Drosophila pseudoobscura*
moderate	151	PMF	6.62	25617	RFW0662_025617	54	0.64	PREDICTED: similar to Sterile alpha motif domain-containing protein 4	*Tribolium castaneum*
moderate	152	PMF	6.20	25651	RFW0620_025651	53	0.9	OS-D-like protein, OS-D2d	*Megoura viciae*
moderate	153	PMF	6.18	26145	RFW0618_026145	54	0.76	heat shock protein Hsp70Ba	*Drosophila melanogaster*
low	154	PMF	6.12	25451	RFW0612_025451	48	2.9	RE20606p	*Drosophila melanogaster*
good	155	PMF	6.13	24721	RFW0613_024721	62	0.12	GA14234-PA	*Drosophila pseudoobscura*
low	156	PMF	6.17	20909	RFW0617_020909	40	17	40S ribosomal protein S5	*Ixodes pacificus*
low	157	PMF	6.40	18246	RFW0640_018246	28	2.7e+002	hypothetical protein AaeL_AAEL005172	*Aedes aegypti* (Yellowfever mosquito)
high	158	MS/MS	6.30	18089	RFW0630_018089	58	4.6	hexamerin II	*Reticulitermes flavipes*
low	159	PMF	6.64	14485	RFW0664_014485	33	85	Drosophila orena nullo	*Drosophila orena*
good	160	PMF	6.44	8270	RFW0644_008270	61	0.16	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	161	PMF	6.51	7596	RFW0651_007596	51	1.5	PREDICTED: similar to CG33957-PB, isoform B	*Tribolium castaneum*
moderate	162	PMF	6.42	7466	RFW0642_007466	52	1	hypothetical protein AaeL_AAEL005876	*Aedes aegypti* (Yellowfever mosquito)
low	163	PMF	6.90	6303	RFW0690_006303	45	6	TPA: TPA_inf: HDC13243	*Drosophila melanogaster*
moderate	164	PMF	6.87	6285	RFW0687_006285	55	0.5	Arginine kinase (AK)	*Cacinus maenas*
low	165	PMF	6.84	7000	RFW0684_007000	39	24	PREDICTED: similar to CG18528-PA	*Tribolium castaneum*
low	166	PMF	6.84	8215	RFW0584_008215	45	5.2	TPA: TPA_inf: HDC13423	*Drosophila melanogaster*
high	167	MS/MS	7.14	9069	RFW0714_009069	195	9.8e-014	hexamerin I	*Reticulitermes flavipes*
low	168	PMF	7.60	8215	RFW0760_008215	47	3.3	ENSANGP00000030419	*Anopheles gambiae* str. PEST
high	169	MS/MS	8.04	8387	RFW0767_007550	131	2.5e-007	hexamirin I	*Reticulitermes flavipes*
good	170	PMF	8.84	8488	RFW0884_008488	60	0.18	ENSANGP00000017345	*Anopheles gambiae* str. PEST

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
good	171	PMF	9.62	8008	RFW0962_008008	58	0.29	GA12861-PA	*Drosophila pseudoobscura*
low	172	PMF	9.57	8214	RFW0957_008214	49	2.3	ENSANGP00000014517	*Anopheles gambiae* str. PEST
low	173	PMF	7.91	10224	RFW0791_010224	41	13	PREDICTED: similar to Coiled-coil domain-containing protein 53	*Apis mellifera*
good	174	PMF	7.29	11554	RFW0729_011554	63	0.096	Ubiquitin	*Arabidopsis thaliana*
low	175	PMF	7.16	12818	RFW0716_012818	43	9.6	DnaJ subfamily B member 11 precursor, putative	*Aedes aegypti* (Yellowfever mosquito)
moderate	176	PMF	7.26	16382	RFW0726_016382	54	0.73	homogentisate 1,2-dioxygenase	*Aedes aegypti* (Yellowfever mosquito)
low	177	PMF	6.97	18550	RFW0697_018550	49	2.1	nanos homolog	*Chironomus samoensis*
good	178	PMF	6.84	17635	RFW0684_017635	63	0.098	CG18528-PA	*Drosophila melanogaster*
low	179	PMF	6.78	16844	RFW0678_016844	49	2.4	zinc finger protein	*Aedes aegypti* (Yellowfever mosquito)
low	180	MS/MS	6.67	20462	RFW0667_020462	51	23	hypothetical protein	*Streptomyces avermitilis* MA-4680
moderate	181	PMF	6.85	23023	RFW0685_023023	55	0.54	ENSANGP00000015486	*Anopheles gambiae* str. PEST
good	182	PMF	7.95	18524	RFW0795_018524	62	0.11	calponin/transgelin	*Aedes aegypti* (Yellowfever mosquito)
moderate	183	PMF	7.83	15508	RFW0783_015508	54	0.69	Putative arginine kinase	*Homalodisca coagulata* (Glassy-winged sharpshooter)
good	184	PMF	6.36	33803	RFW0636_033803	57	0.32	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	185	PMF	6.30	36043	RFW0630_036043	45	5.7	GA19753-PA	*Drosophila pseudoobscura*
low	186	PMF	6.97	43594	RFW0697_043594	47	3.6	CG11023 protein	*Drosophila melanogaster*
high	187	MS/MS	7.34	40554	RFW0734_040554	164	1.2e-010	Cuticle protein 21 (LM-ACP 21)	*Locusta migratoria*
high	188	MS/MS	7.36	34211	RFW0736_034211	297	6.2e-024	beta-actin	*Spodoptera littoralis*
high	189	MS/MS	7.43	33059	RFW0743_033059	166	7.8e-011	PREDICTED: similar to heat shock protein hsp21.4	*Apis mellifera*
good	190	PMF	9.63	17770	RFW0963_017770	56	0.43	PREDICTED: similar to CG32158-PB, isoform B	*Apis mellifera*
good	191	PMF	9.31	16871	RFW0931_016871	58	0.3	hypothetical protein AaeL_AAEL003229	*Aedes aegypti* (Yellowfever mosquito)
good	192	PMF	9.35	14594	RFW0935_014594	57	0.37	LD40801p	*Drosophila melanogaster*
low	193	PMF	4.30	38056	RFW0430_038056	41	15	PREDICTED: similar to CG14672-PA, isoform 1	*Tribolium castaneum*
low	194	PMF	4.22	42676	RFW0422_042676	47	3.3	unnamed protein product	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	195	PMF	4.38	33683	RFW0438_033683	37	39	iPLA-1	*Aedes aegypti* (Yellowfever mosquito)
low	196	PMF	5.30	50084	RFW0530_050084	39	24	CG7603-PA	*Drosophila melanogaster*
low	200		3.42	17442	RFW0342_017442				
good	202	PMF	4.53	11862	RFW0453_011862	57	0.32	short stop CG18076-PC, isoform C	*Drosophila melanogaster*
moderate	204	PMF	5.95	18590	RFW0595_018590	53	0.94	GA16441-PA	*Drosophila pseudoobscura*
low	207	PMF	6.00	28929	RFW0600_028929	28	3e+002	PREDICTED: similar to CG12926-PA isoform 1	*Tribolium castaneum*
low	208	PMF	6.09	27416	RFW0609_027416	51	1.3	Luciferase	*Rhagophthalmus ohbai*
high	209	MS/MS	9.71	33764	RFW0971_033764	207	6.2e-015	beta-actin	*Bubalus bubalis*
good	210	PMF	9.63	33718	RFW0963_033718	61	0.15	ENSANGP00000030634	*Anopheles gambiae* str. PEST
moderate	211	PMF	7.09	33382	RFW0709_033382	55	0.57	acyl-coa dehydrogenase	*Aedes aegypti* (Yellowfever mosquito)
moderate	212	PMF	7.02	35105	RFW0702_035105	53	0.95	PREDICTED: similar to CG5890-PA	*Apis mellifera*
low	213	PMF	6.93	32140	RFW0693_032140	47	3.5	ribosomal protein L35Ae	*Sphaerius* sp. APV-2005
moderate	214	PMF	6.81	27015	RFW0681_027015	55	0.55	Protein on ecdysone puffs CG6143-PB, isoform B	*Drosophila melanogaster*
high	215	MS/MS	5.99	27069	RFW0599_027069	162	2e-010	PREDICTED: similar to tropomyosin 1	*Apis mellifera*
low	216	PMF	5.68	24570	RFW0568_024570	31	1.3e+002	ENSANGP00000016550	*Anopheles gambiae* str. PEST
low	217	PMF	5.74	21583	RFW0574_021583	48	2.8	Cytochrome P450-9b2 CG4486-PA	*Drosophila melanogaster*
high	218	MS/MS	6.24	26767	RFW0624_026767	129	3.9e-007	beta-tubulin	*Salpingoeca amphoridium*
low	219	PMF	6.71	22934	RFW0671_002934	45	6.1	6-phophogluconate dehydrogenase	*Aedes aegypti* (Yellowfever mosquito)
high	220	PMF	7.93	12907	RFW0793_012907	65	0.059	CG6630-PA	*Drosophila melanogaster*
low	221	PMF	8.45	12627	RFW0845_012627	37	36	CG11251-PA	*Drosophila melanogaster*
good	222	PMF	8.55	12625	RFW0855_012625	64	0.066	Muscle-specific protein 300 CG33715-PB, isoform B	*Drosophila melanogaster*
low	223	PMF	8.63	10918	RFW0863_010918	38	29	ENSANGP00000016055	*Anopheles gambiae* str. PEST
good	224	PMF	8.45	29940	RFW0845_029940	60	0.16	elongation factor 1 alpha	*Battus polydamas*
high	225	MS/MS	8.65	27104	RFW0865_027104	63	1.6	hexamerin II	*Reticulitermes flavipes*
moderate	227	PMF	9.54	16679	RFW0954_016679	54	0.7	elongation factor-1 alpha	*Euryglossa ephippiata*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
high	228	MS/MS	9.61	14518	RFW0961_014518	141	2.5e-008	Superoxide dismutase [Cu-Zn]	*Ceratitis capitata*
good	229	PMF	9.63	15030	RFW0963_015030	58	0.26	ENSANGP00000030350	*Anopheles gambiae* str. PEST
low	230	PMF	9.36	16435	RFW0936_016435	43	8.6	ENSANGP00000021758	*Anopheles gambiae* str. PEST
moderate	231	PMF	9.28	13170	RFW0928_013170	54	0.71	mitochondrial carrier protein	*Aedes aegypti* (Yellowfever mosquito)
low	233	PMF	9.54	13823	RFW0954_013823	52	1	PREDICTED: similar to CG17947-PA	*Tribolium castaneum*
moderate	234	PMF	9.35	26747	RFW0935_026747	53	0.96	triosephosphate isomerase	*Calliphora vicina* (Blue blowfly, Calliphora erythrocephala)
low	236	PMF	6.86	14473	RFW0686_014473	41	16	AT25667p	*Drosophila melanogaster*
low	237	PMF	7.05	11794	RFW0705_011794	45	5.1	ENSANGP00000003286	*Anopheles gambiae* str. PEST
high	238	MS/MS	8.83	6724	RFW0883_006724	174	1.2e-011	hexamerin I	*Reticulitermes flavipes*
high	239	MS/MS	8.96	16832	RFW0896_016832	89	0.0037	ENSANGP00000003966	*Anopheles gamiae* str. PEST
low	239	CAF	8.96	16832	RFW0896_016832				
low	240	PMF	7.17	11658	RFW0717_011658	50	1.9	Twist	*Achaearanea tepidariorum*
low	241	PMF	8.34	29611	RFW0834_029611	31	1.4e+002	GA11134-PA	*Drosophila pseudoobscura*
moderate	242	PMF	8.54	26664	RFW0854_026664	53	0.96	CG31699-PA	*Drosophila melanogaster*
low	244	PMF	6.82	12745	RFW0682_012745	48	3	elongation factor 1 alpha	*Xyleborus sphenos*
good	245	PMF	6.83	11854	RFW0683_011854	57	0.36	Putative arginine kinase	*Homolodisca coagulata*
low	246	PMF	7.00	11230	RFW0700_011230	47	3.5	PREDICTED: similar to CG9422-PA, isoform A	*Tribolium castaneum*
low	247	PMF	6.34	8771	RFW0634_008771	46	4.1	PREDICTED:similar to Profilin (Chickadee protein)	*Tribolium castaneum*
low	248		6.28	6097	RFW0628_006097				
low	249	PMF	5.97	5726	RFW0597_005726	42	11	ENSANGP00000025197	*Anopheles gambiae* str. PEST
low	250		5.59	6440	RFW0559_006440				
low	251	PMF	6.65	5927	RFW0665_006927	48	2.7	2-oxoglutarate dehydrogenase	*Aedes aegypti* (Yellowfever mosquito)
low	252	PMF	7.01	6153	RFW0701_006153	46	4.1	PREDICTED: similar to CG3295-PA	*Apis mellifera*
low	253	PMF	7.14	10551	RFW0714_010551	41	14	malic enzyme	*Aedes aegypti* (Yellowfever mosquito)
low	255	PMF	7.61	10115	RFW0761_010115	50	1.8	PREDICTED: similar to Cytochrome P450 family member (cyp-31A3)	*Tribolium castaneum*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
good	256	PMF	6.72	36296	RFW0672_036296	56	0.42	26S proteasome subunit	*Aedes aegypti* (Yellowfever mosquito)
low	257	PMF	7.35	43835	RFW0735_043835	46	4.2	CG33095-PA, isoform A	*Drosophila melanogaster*
high	258	PMF	7.86	17151	RFW0786_017151	68	0.03	CG15611-PB, isoform B	*Drosophila melanogaster*
low	259	PMF	9.56	6317	RFW0956_006317	41	15	sniffer CG0964-PA	*Drosophila melanogaster*
low	260	PMF	9.34	5439	RFW0934_005439	51	1.3	Heat shock protein 60 related CG2830-PA	*Drosophila melanogaster*
low	261	PMF	4.46	5858	RFW0446_005858	29	2.3e+002	RNA polymerase II largest subunit [parasitid 'Pas']	*parasitid* 'Pas'
low	262	PMF	4.40	5958	RFW0440_005958	31	1.4e+002	CG41090-PA.3	*Drosophila melanogaster*
low	263	PMF	4.34	6886	RFW0434_006886	47	3.3	PREDICTED: similar to CG5991-PA, isoform A	*Tribolium castaneum*
low	343	PMF	4.18	65459	RFW0418_065459	22	1.2e+003	CG15375-PA	*Drosophila melanogaster*
low	345		4.21	67227	RFW0421_067227				
low	347		4.19	71363	RFW0419_071363				
good	408	PMF	4.60	16618	RFW0460_016618	64	0.071	Ga19599-PA	*Drosophila pseudoobscura*
low	446	PMF	4.50	17894	RFW0450_017894	42	11	ENSANGP00000002307	*Anopheles gambiae* str. PEST
high	483	MS/MS	5.38	41754	RFW0538_041754	211	2.50E-15	beta-actin	*Bubalus bubalis*
good	485	PMF	5.48	39605	RFW0548_039605	56	0.49	GA13812-PA	*Drosophila pseudoobscura*
low	486	PMF	5.62	39999	RFW0562_039999	41	13	RNA polymerase II largest subunit	*Lithobius forficatus*
low	500	PMF	5.37	57577	RFW0537_057577	42	11	PREDICTED: similar to guanine nucleotide binding protein-like 1	*Apis mellifera*
low	505	PMF	5.38	53927	RFW0538_053927	50	1.9	GA17278-PA	*Drosophila pseudoobscura*
low	506	PMF	5.34	52029	RFW0534_052029	51	1.5	12kDa hemolymph protein d precursor	*Tenebrio molitor*
high	508	MS/MS	5.30	45384	RFW0530_045384	98	0.00047	Actin-5, muscle specific	*Bactrocera dorsalis*
high	509	PMF	5.40	40427	RFW0540_040427	68	0.03	muscle-specific calpain	*Gecarcinus laterealis*
low	569	PMF	4.78	28485	RFW0478_027485	51	1.3	GA17307-PA	*Drosophila melanogaster*
moderate	570	PMF	4.79	26268	RFW0479_026268	55	0.6	ENSANGP00000023972	*Anopheles gambiae* str. PEST
low	573	PMF	4.87	32871	RFW0487_032871	42	12	RBF	*Drosophila melanogaster*
good	574	PMF	4.88	35826	RFW0488_035826	57	0.39	GA21222-PA	*Drosophila pseudoobscura*
good	575	PMF	4.84	38094	RFW0484_038094	57	0.33	Rab9 CG9994-PA	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	596	PMF	6.08	47398	RFW0608_047398	38	30	PREDICTED: similar to Mitochondrial import inner membrane translocase subunit Tim8	*Apis mellifera*
low	610	PMF	6.00	25980	RFW0600_025980	40	16	PREDICTED: similart to photoreceptor-specific nuclear receptor isoform B	*Tribolium castaneum*
low	622	PMF	6.25	49813	RFW0625_049813	51	1.5	GA15903-PA	*Drosophila pseudoobscura*
low	626	PMF	5.63	36457	RFW0563_036457	42	12	PREDICTED: similar to 4GT2 CG5878-PA, partial	*Apis mellifera*
high	635	PMF	6.03	29396	RFW0603_029396	86	0.00045	putative secreted salivary protein	*Ixodes scapularis*
low	640	PMF	6.20	31750	RFW0620_031750	42	12	stathmin CG31641-PA, isoform A	*Anopheles gambiae* str. PEST
low	642	PMF	6.19	32963	RFW0619_032963	41	14	PREDICTED: similar to CG7528-PA	*Tribolium castaneum*
low	645	PMF	6.58	30076	RFW0658_030076	39	20	CG6067-PA	*Drosophila melanogaster*
low	647	PMF	6.32	28359	RFW0632_028359	50	1.7	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	650	PMF	6.59	27964	RFW0659_027964	35	52	ENSANGP00000011027	*Anopheles gambiae* str. PEST
good	654	PMF	7.15	28309	RFW0715_028309	57	0.36	CG9897	*Drosophila melanogaster*
low	685	PMF	8.83	13297	RFW0883_013297	45	5.7	PREDICTED: similar to activating transcription factor 2 isoform 1	*Apis mellifera*
high	686	MS/MS	9.08	12757	RFW0908_012757	93	0.0017	muscular protein 20	*Cicindela nimuta*
low	694	PMF	9.93	12102	RFW0993_012102	52	1.2	PREDICTED: similar to chromosome condensation protein G	*Tribolium castaneum*
low	695	PMF	9.91	11307	RFW0991_011307	48	2.8	DEAD-box protein	*Drosophila melanogaster*
high	698	PMF	9.51	10401	RFW0951_010401	69	0.025	GH03748p	*Drosophila melanogaster*
low	714	PMF	6.79	10191	RFW0679_010191	47	3.7	CG18528-PA	*Drosophila melanogaster*
low	717	PMF	7.22	8463	RFW0722_008463	37	35	TPA: TPA_inf: HDC12873	*Drosophila melanogaster*
low	720	PMF	6.80	7374	RFW0680_007374	48	2.9	PREDICTED: similar to dynactin 2 (p50)	*Apis mellifera*
low	726		6.04	7048	RFW0604_007048				
low	737		6.25	15214	RFW0625_015214				
low	739	PMF	6.41	15533	RFW0641_015533	39	20	GA10385-PA	*Drosophila pseudoobscura*
low	740	PMF	6.36	16375	RFW0636_016375	44	6.4	ATM protein	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Exp. pI	Exp. MW	Protein Identifier	Score	Expect	Description	Source
low	748	PMF	6.18	19269	RFW0618_019269	44	7.7	CG9617	*Drosophila yakuba* (Fruit fly)
low	749	PMF	6.11	21036	RFW0611_021036	38	25	OS-D-like protein, OS-D2a	*Metopolophium dirhodum*
low	750	PMF	6.12	22236	RFW0612_022236	33	83	CG17601-PA	*Drosophila melanogaster*
low	751	PMF	6.11	22805	RFW0611_022805	43	8.9	phosphatidylethanolamine-binding protein	*Aedes aegypti* (Yellowfever mosquito)
low	752	PMF	6.18	24126	RFW0618_024126	50	1.7	malic enzyme	*Aedes aegypti* (Yellowfever mosquito)
low	753	PMF	6.18	23085	RFW0618_023085	47	3.8	GA21222-PA	*Drosophila pseudoobscura*
high	754	PMF	6.25	21021	RFW0625_021021	65	0.05	RHC18, putative	*Aedes aegypti* (Yellowfever mosquito)
good	764	PMF	5.11	21141	RFW0511_021141	58	0.25	PREDICTED: similar to centrosome protein cep290	*Tribolium castaneum*
good	771	PMF	5.61	15683	RFW0561_015683	58	0.3	cytochrome P450 CYP6P8	*Anopheles minimus*
moderate	776	PMF	4.78	12436	RFW0478_012436	55	0.54	mus81 CG3026-PA	*Drosophila melanogaster*
good	778	PMF	4.49	13182	RFW0449_013182	57	0.37	CG11414-PA	*Drosophila melanogaster*
low	784	PMF	4.91	6037	RFW0491_006037	43	8.7	ENSANGP00000030805	*Anopheles gambiae* str. PEST
moderate	789	MS/MS	4.30	15106	RFW0430_015106	59	4.2	putative protein	*Arabidopsis thaliana*
high	790	PMF	4.11	19281	RFW0411_019281	67	0.035	GA198964-PA	*Drosophila pseudoobscura*
moderate	791	PMF	4.15	19892	RFW0415_019892	55	0.61	Tes154	*Drosophila mojavensis* (Fruit fly)
low	797	PMF	4.34	23389	RFW6434_023389	49	2.2	ENSANGP00000011312	*Anopheles gambiae* str. PEST
moderate	798	PMF	6.48	73947	RFW0648_073947	54	0.73	technical knockout CG7925-PB	*Drosophila melanogaster*
moderate	800	PMF	5.65	35360	RFW0565_035360	55	0.61	ENSANGP00000029084	*Anopheles gambiae* str. PEST
low	804	PMF	3.96	82270	RFW0396_082270	31	1.5e+002	microtubule-associated protein	*Aedes aegypti* (Yellowfever mosquito)

End of Table

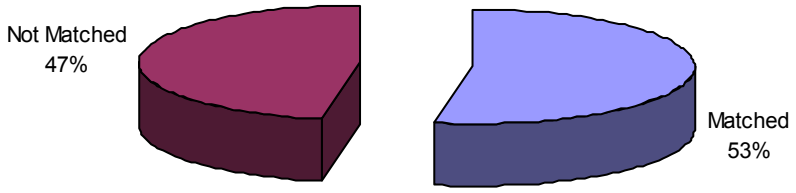


Fig. 25. Percent proteins matched and not matched to molecular function.

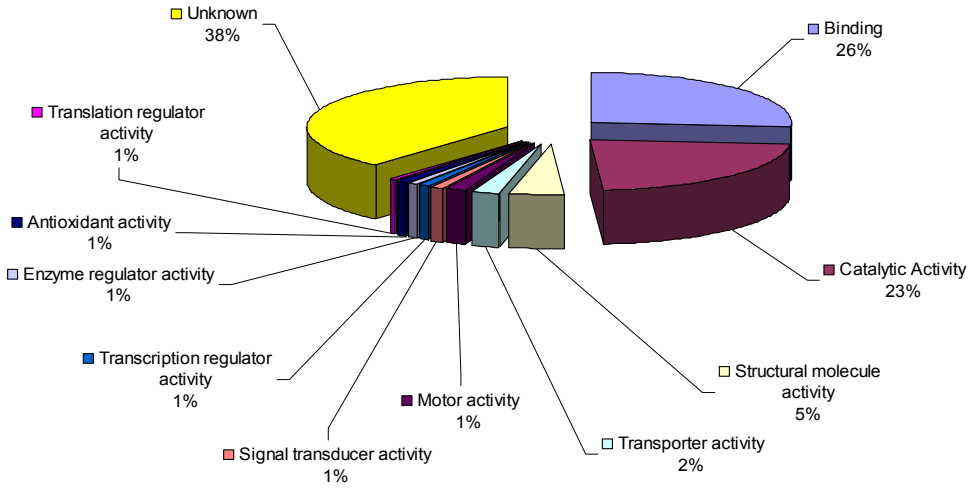


Fig. 26. Distribution of molecular functions.

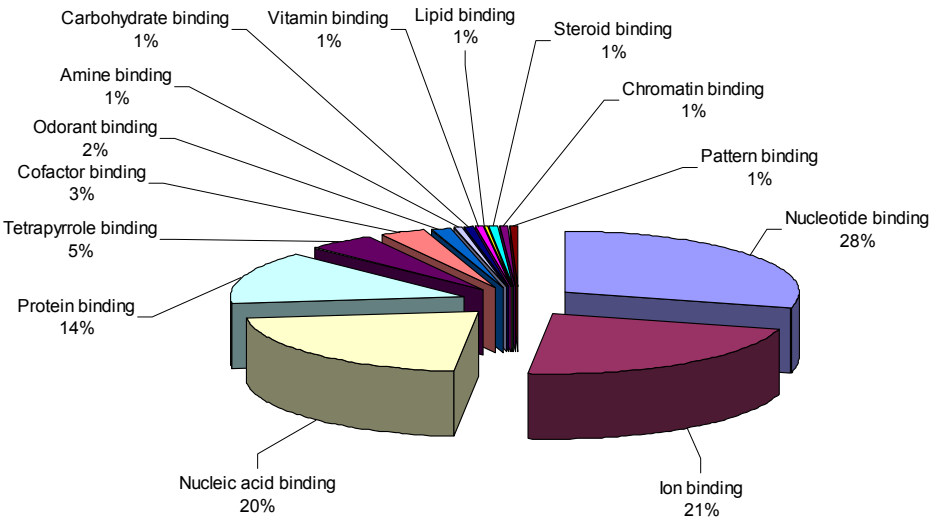


Fig. 27. Distribution of binding functions.

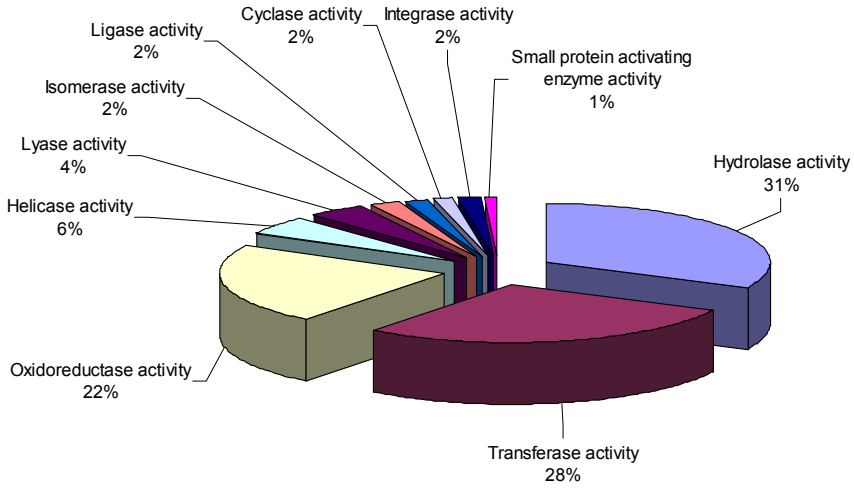


Fig. 28. Distribution of catalytic activity.

Objective IV – Test for differential protein expression among *R. flavipes* colonies

Multiple two-dimensional gels were generated for each colony. Gels were compared to test for differentially expressed proteins among the colonies. Analysis was undertaken using both manual comparison and gel analysis software. Differential expression can appear in several forms. It can be expressed as up or down regulated proteins where proteins are present in higher concentrations (up regulated) or lower concentrations (down regulated). This condition can be presented as slight changes in spot density or complete absence/presence of spots. It is possible visibly absent proteins may be present, but in concentrations below the threshold of detectibility.

Dimension software was used to test for differential protein expression among *Reticulitermes flavipes* colonies. Several gel analysis programs were evaluated. Each program was capable of fulfilling the basic criteria of identifying protein spots, warping gels to align spots, and comparing common proteins among samples. However, each program varied with its ability to intuitively interface with the user. Two programs were selected for final comparison, ImageMaster™ 2D and Dymension. Dymension was selected since it provided a more user-friendly interface and facilitated use by a casual user. Further consideration was the selection of 'ImageMaster™ 2D Platinum DIGE Enabled' by the Biochemistry and Molecular Biochemistry Core Facility, allowing eventual access to both of the final selections.

Dymension software was best suited to comparing common proteins among samples, but did provide capacity for observing differentially expressed proteins. Dymension provided no manner of comparison for proteins considered non-consensus.

When non-consensus spots were identified that should have been marked as consensus, spots could be marked manually on gels to facilitate their detection as a consensus spot.

Dymension normalized protein spot densities based on the master gel. On the master gel, the density for every spot was one. As additional gels were compared, the spot densities were quantitatively estimated relative to the master gel. Comparison of these estimates allowed elucidation of up-regulated or down-regulated spots among samples. Another method of searching for differentially expressed proteins was overlaying the colony sample images and visually comparing spot outlines (Fig. 29). However, as samples were added to the overlay, the spot outlines rapidly became difficult to compare. Again, this method was more useful in comparing similarities rather than differential protein expression.

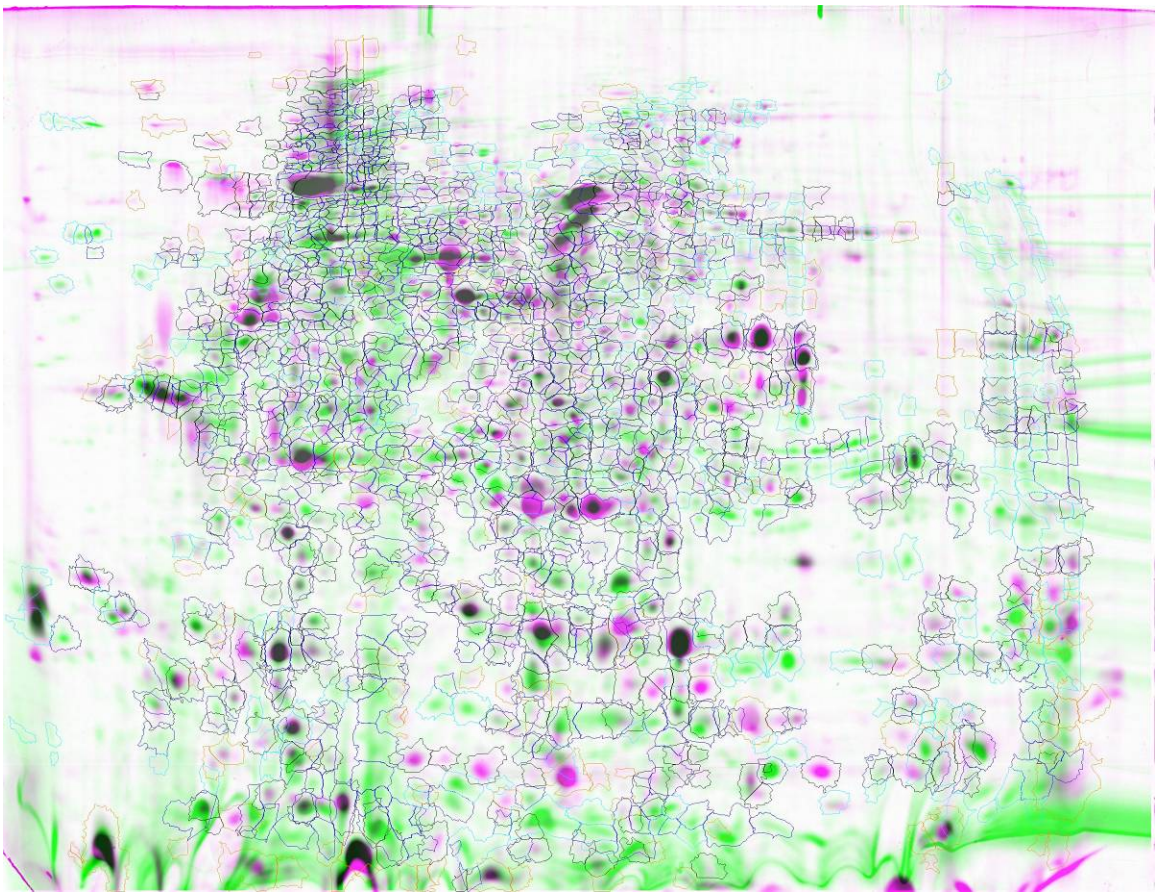


Fig. 29. Overlayed Dymension gel images and spot outlines.

Samples were generated using gel images from three replicates from each colony. Samples 1, 2, and 3 corresponded to Colony 1, 2, and 3, respectively. The Colony 1 sample established the basic map to which the other colony samples were compared and consensus spots were outlined (Fig. 30). Each additional colony sample was compared to the first sample and a table of consensus spots was assembled.

Preliminary gel comparison during protocol establishment suggested matching gel maps among colonies. Closer inspection revealed variation in the spot patterns among the replicates. Comparison of the Colony 1 replicates (Fig. 31) yielded 909 consensus spots, Colony 2 replicates (Fig. 32) yielded 988 consensus spots, and Colony 3 replicates (Fig. 33) yielded 960 consensus spots. There was a difference of seventy-nine consensus spots between Colony 1 and Colony 2, twenty-eight consensus spots between Colony 2 and Colony 3, and fifty-one consensus spots between Colony 1 and Colony 3. As the Colony 1 sample was collected on the OSU campus at Stillwater, OK and the Colony 2 and Colony 3 samples were collected on the Tallgrass prairie near Pawhuska, OK, this indicates a noteworthy difference between the two geographic regions. Differentially expressed proteins accounted for less than ten percent of the total spots yielded. Although differences were observed, the protein maps generated from each colony had a substantial number of common proteins. Approximately 50 spots were well defined, and had adequate density among gels to make them useful as markers for gel warping to minimize spatial variations among protein spot patterns. Reproducible markers indicated the protein maps were similar and may be useful as a taxonomic tool. However, additional studies with other termite species will be necessary to validate any potential taxonomic value.

Manual comparison of gels was used to confirm Dymension results. A gel image from each colony was selected and divided into three equal sections. Sections were recombined with the corresponding sections from the other images to create sectional gel images (Figs. 34–36). While images such as these facilitate multi-gel comparison, they also illustrate the difficulty of matching spots without the advantage of gel warping technology. Differentially expressed proteins could be found in various combinations of the three images. For example, Fig. 34A, E, and G; Fig. 35I, and Fig. 36Y demonstrate proteins that are present in Colony 1 and Colony 2, but absent in Colony 3. Proteins present in Colony 2 and Colony 3, but absent in Colony 1 are represented by Fig. 34F, Fig. 35N and O, and Fig. 36T and V. Proteins present in Colony 1 and Colony 3, but absent in Colony 2 are represented by Fig. 35M, Fig. 36P, U, X, and Z. Proteins specified by Fig. 35K and L are only found on the Colony 1 gel, while proteins specified by Fig. 34C and Fig. 36R are only found on the Colony 3 gel. At the other end of possible combinations, are proteins present in each colony, but variations in gel separation make it difficult to accurately assess the grouping as illustrated in Fig. 34B and D, Fig. 35H and J, and Fig. 36S, Q, and W. Non-separated groups of proteins were treated as a single spot to facilitate analysis and due to the necessity of target these proteins specifically using small pH range IPG strips and different acrylamide concentrations. These examples emphasize why digital gel warping and comparison software are rapidly becoming critical components of two-dimensional gel analysis.

Although the primary goal of establishing an overall protein profile was achieved, differential protein analysis was complicated by the limitation of using a single pI range of pH 3–10. Improved analysis would have been achieved by selecting multiple smaller

pH ranges and acrylamide concentrations. This would have created multiple gel maps that could be combined to generate a more accurate overall image. Future studies should consider having a minimum of three pH ranges and two concentrations yielding at least six sections. This would allow increased resolution and greater protein separation facilitating better protein matching and identification. One caveat to the increased number of sections would be the necessity of greater sample requirements. For example, a gel having a 4–7 pH range would require approximately twice the amount of protein as a gel having a 3–10 pH range and a gel having a 5–6 pH range would require approximately four times the protein to generate an equivalent spot. In other words, the smaller the pH range, the more protein is required to achieve the equivalent result.

Two-dimensional gel electrophoresis is an extremely capable method for comparing large quantities of proteins. However, this method is not without its flaws. For example, gel-to-gel variation can cause difficulties in aligning spot patterns for comparison. Additionally, gels completed using protein from the same can exhibit variations in density. Currently, the best method to perform differential protein analysis is fluorescence two-dimensional difference gel electrophoresis (DIGE). DIGE allows multiple samples to be separated simultaneously on the same gel. Typically, this technique is completed with a control and two samples, each labeled with a fluorescent dye visible at a different wavelength. The gel can be scanned and the images merged to compare differences between the samples. This could help eliminate variation between samples, but variation among replicates would remain. Variation among gels is remedied by using an internal standard comprised of all the samples, thus providing a consistent

method to normalize spot densities among gels. The BMB core facility is in the process of acquiring this equipment and it should be available for future studies.

Confounding issues – Speculation of causes for differentially expressed proteins among termite colonies yields several possibilities. The most obvious possibility would be differences in the protein profile due to the geographic region. Causes for consideration could be slight variance in similar proteins due to food sources or adaptations to environmental stresses such as variations in temperature range or water availability. This would be tentatively supported by the higher number of differentially expressed proteins from the Tallgrass Prairie termite colony samples relative to the Stillwater sample. Another geographically influenced possibility may be differences in the symbiont profiles. As a whole body extract was used, it is possible protist and bacteria proteins may be present in the protein profile. If the symbionts varied sufficiently, it is possible differences may have presented as differentially expressed proteins.

Other explanations may apply. For example, this study did not differentiate between worker instars during specimen collection as multiple worker instars were combined to form the worker sample. Since immature termites are considered workers during the third instar and beyond, it would be possible for developing termites to have variations in their protein profile. Thus, some of differentially expressed proteins may be due to variations in the age ranges of the termites collected in a sample. This would be tentatively supported by the differing ages of the lab 'colonies'. As termites are collected in the wild, and moved to a laboratory setting, colony dynamics are changed from those of the field colony, and the collected termites will begin to establish a new colony. Thus,

some workers may develop into secondary reproductives and initiate reproduction. This will begin to stabilize the caste dynamics, but will also introduce early instar immature termites into the laboratory colony. As the Stillwater 'colony' is the oldest, it is possible that caste dynamics may have affected the protein profiles. However, each colony is purposely kept in the lab for a period of time to stabilize and to minimize outside influences. This stablilization is necessary to normalize each colony to equivalent conditions. If equivalent conditions exist, additional studies may reveal some differential expressed proteins between instars. We already know significant differences occur between worker and soldier castes (see Objective V).

A third explanation of differentially expressed proteins could be seasonal effects on the caste dynamics. Termites respond to seasonal and/or environmental cues for colony activities such as swarming. In the wild, alates generally swarm on warm spring days with high humidity. There is seasonal variation among different species of termites. However, captive termites are kept in darkened containers in temperature control laboratories. This removes many of the seasonal cues that may influence shifts in caste dynamics of wild colonies, such as the production of alates for spring swarms. Although laboratory colonies are typically small, alate production and swarming activity was observed in the laboratory. These swarms occur during the correct season and conditions independent of many obvious influences. Since the termites for each replicate were not collected during the same season, seasonal influences may have contributed to the variation in protein profiles among colonies. This potential influence could be removed from future studies by processing enough termites during a single collection to generate adequate precipitant for each replicate in the study.

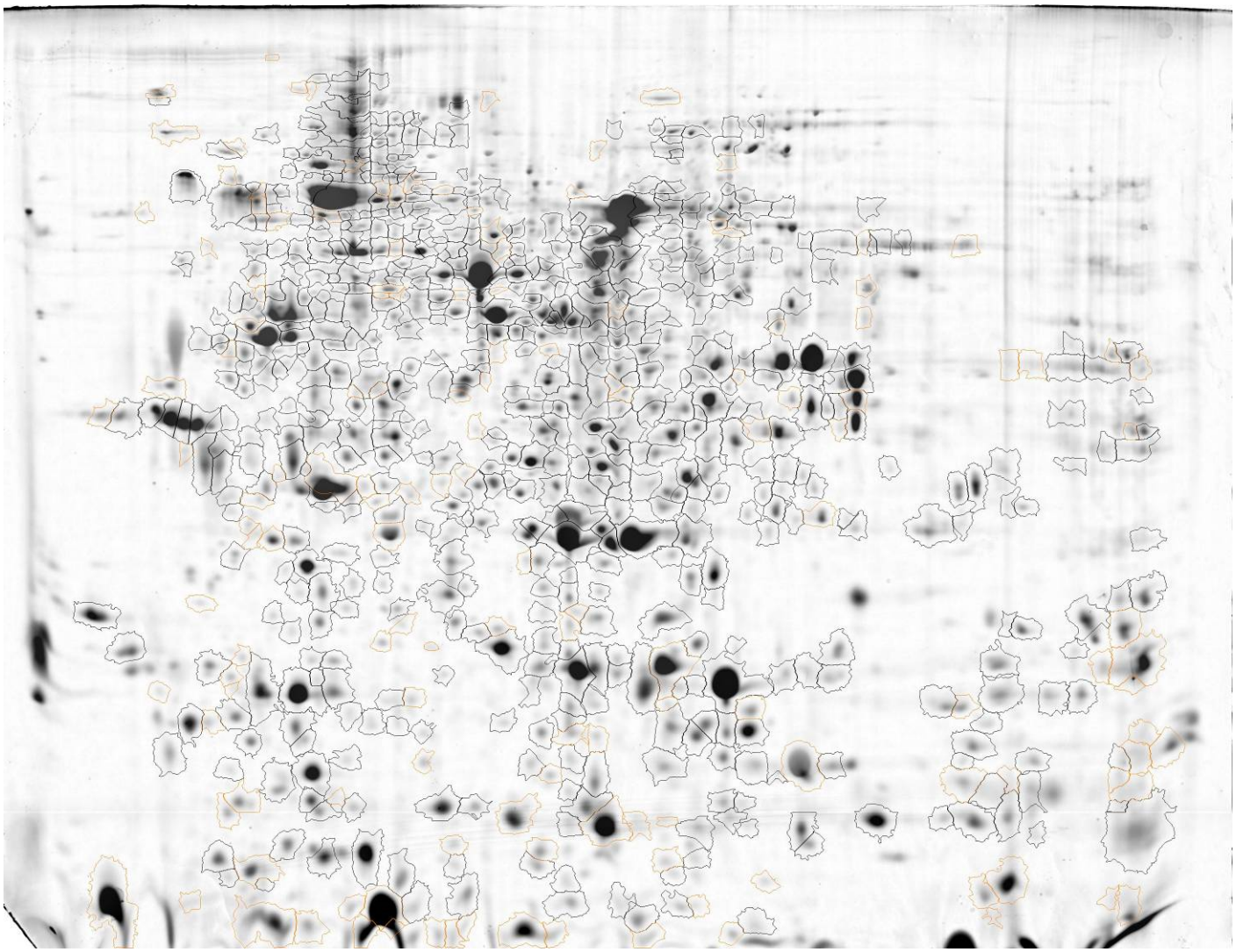


Fig. 30. Dymension image for consensus spots common among samples.

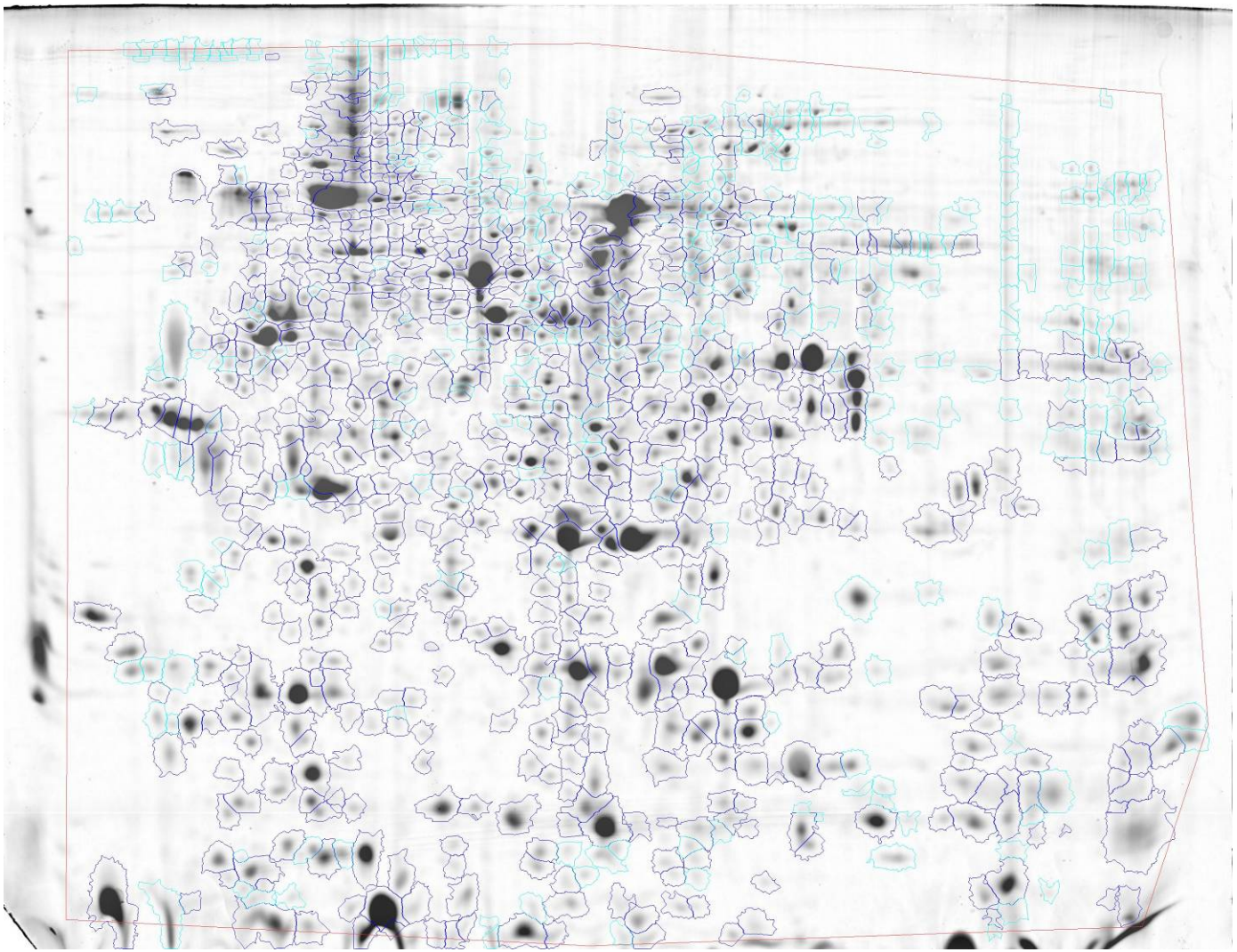


Fig. 31. Dymension image for 'Colony 1' consensus spots.

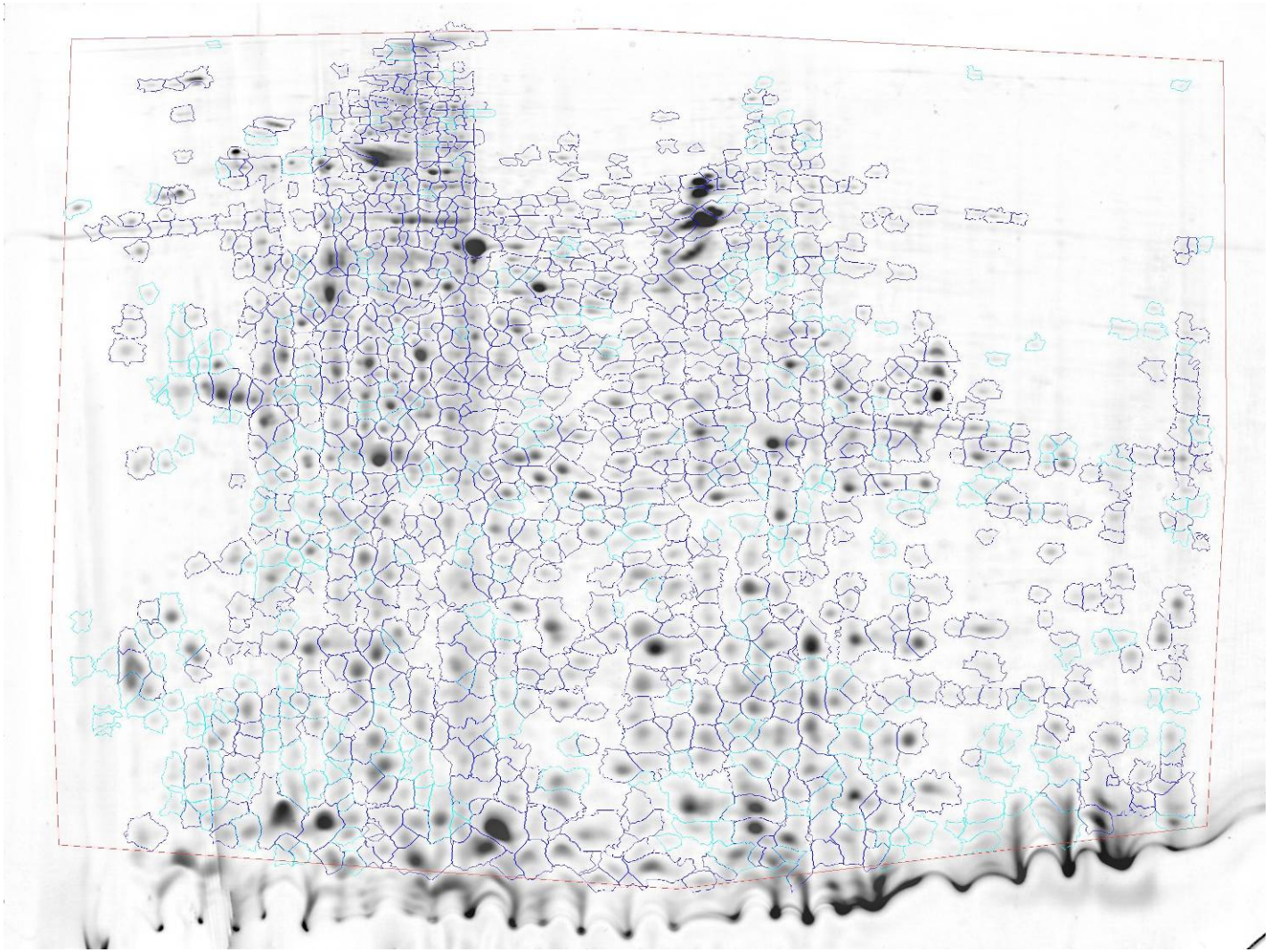


Fig. 32. Dymension image for 'Colony 2' consensus spots.

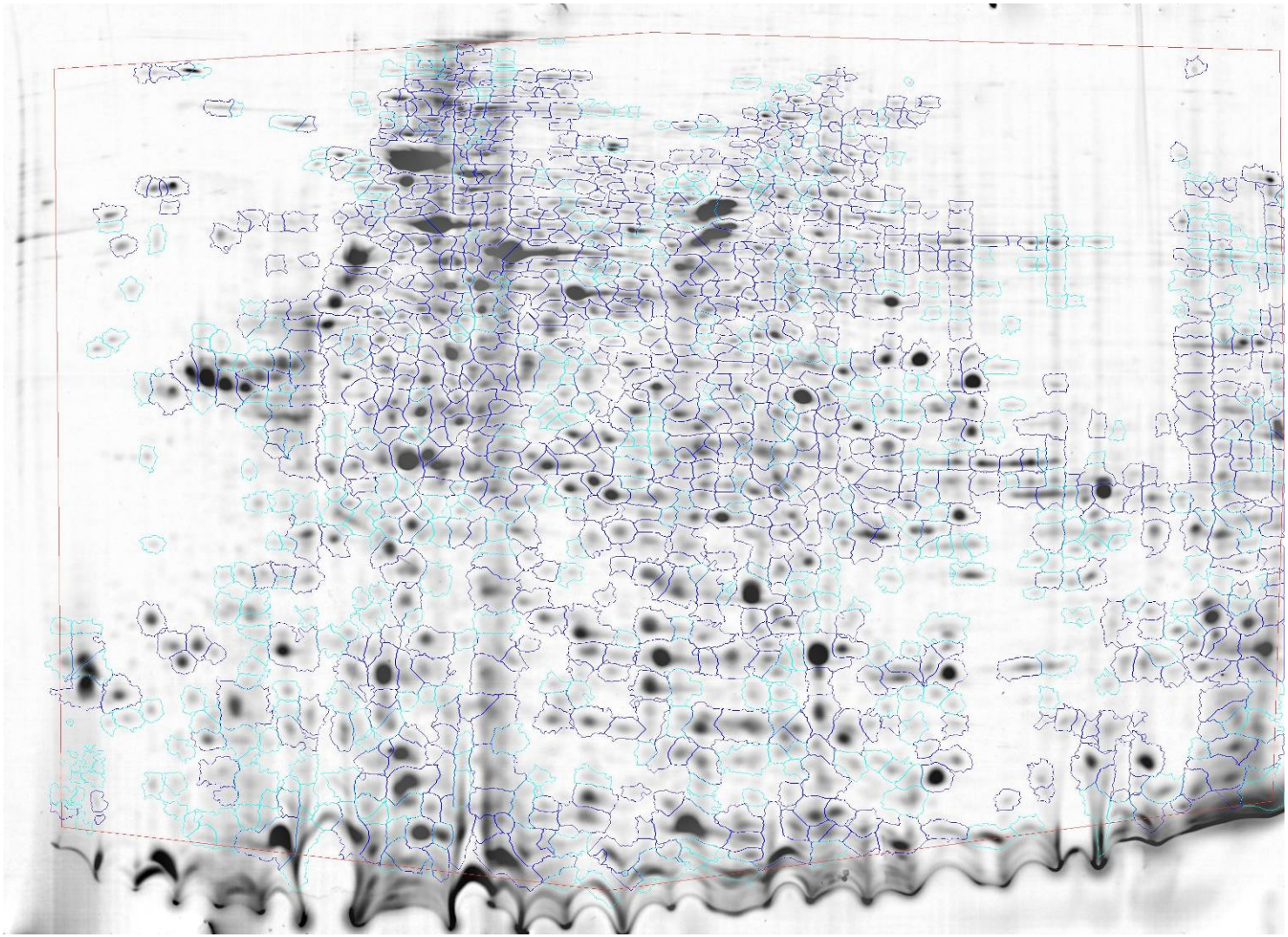


Fig. 33. Dymension image for ‘Colony 3’ consensus spots.

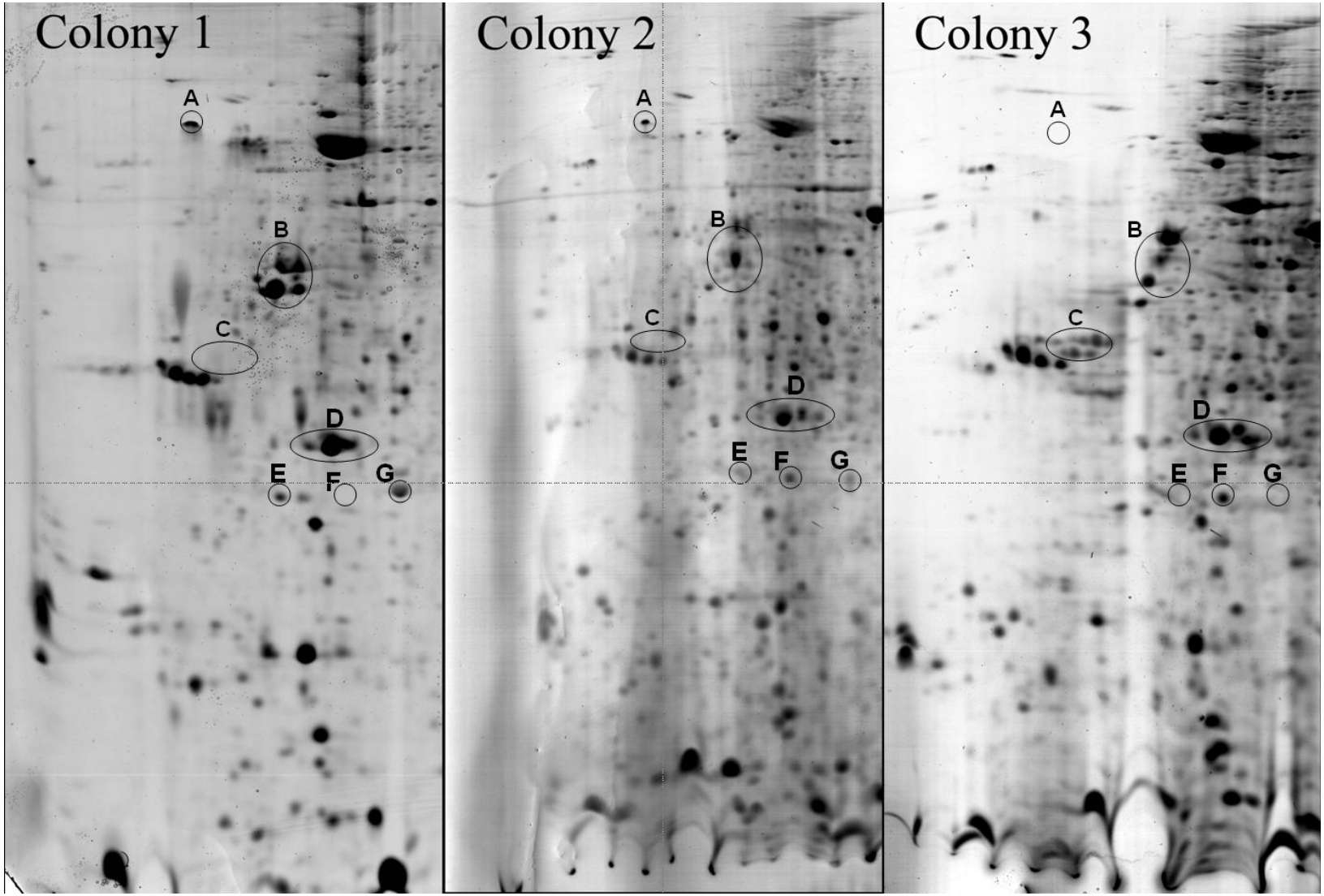


Fig. 34. Image for colony comparison – Section 1.

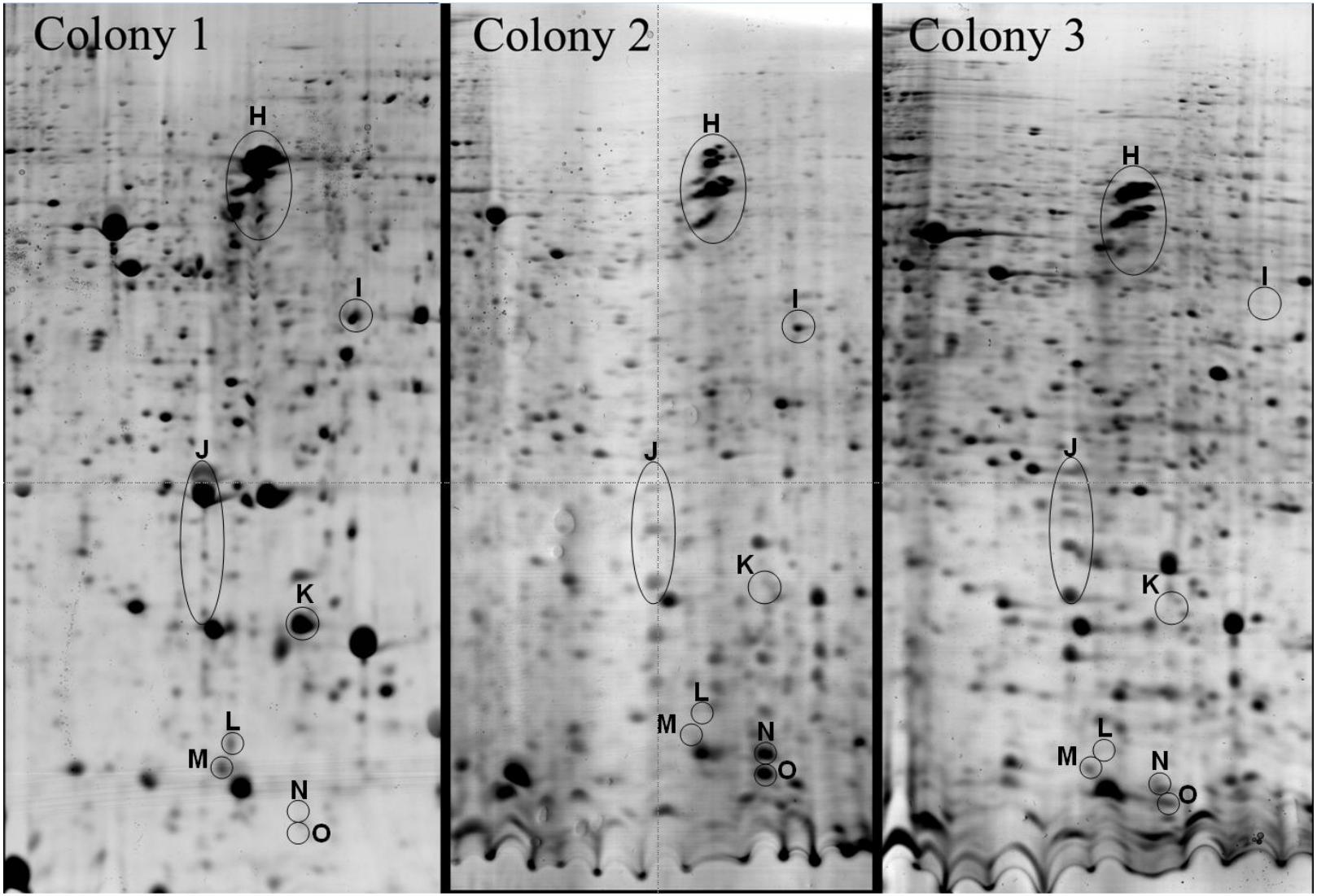


Fig. 35. Image for colony comparison – Section 2.

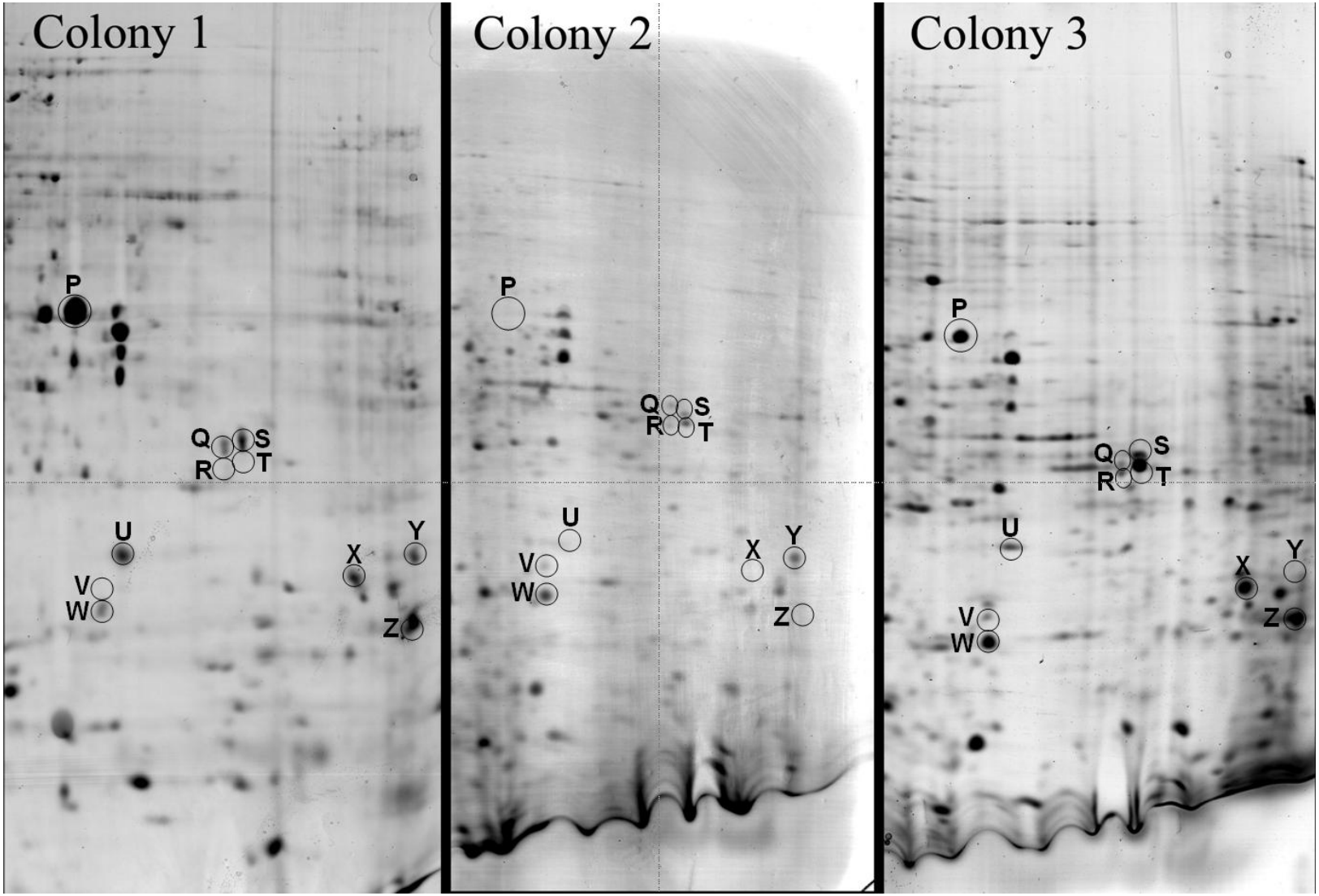


Fig. 36. Image for colony comparison – Section 3.

Objective V – Test for differential protein expression between worker and soldier castes

A reference map for the soldier caste of Colony 1 (Fig. 37) was generated as described in Materials and Methods, Objective 2 – Image analysis, and divided into quadrants (Figs. 38–41) to facilitate viewing. The reference maps for worker and soldier castes allowed the differential comparison between the worker caste and soldier caste of Colony 1. Colony 1 gels from the colony comparison were used since the soldiers were also collected from Colony 1 during the worker collections. However, due to the ratio of soldiers-to-workers, approximately five soldiers per one hundred workers, soldiers were collected during each worker collection, frozen, and stored at -80°C until adequate specimens were collected for a sample. Since soldiers are a terminal-form, potential seasonal variations in caste dynamics affecting workers should have minimal impact.

Dymension software was used to observe differential comparison between the worker and soldier castes (Fig. 42) using samples comprised of gel images from each caste. While Dymension software was proficient at identifying and displaying common proteins among multiple samples, it was not as efficient for displaying differentially expressed proteins. However, Dymension did prove to be a valuable tool for elucidating differentially expressed proteins.

A Dymension analysis comparison of the Colony 1 worker caste replicates (Fig. 43) yielded 855 consensus spots. Dymension comparison of the Colony 1 soldier caste replicates (Fig. 44) yielded 837 consensus spots. However, comparing the spot maps of the worker and soldier castes yielded only 423 common proteins between the consensus spots. This would imply 432 differentially expressed proteins in the worker

caste and 414 differentially expressed proteins in the soldier caste. The high number of non-matched consensus spots indicates substantial difference between the protein profiles of worker and soldier castes.

Dymension provided an efficient method for estimating pI and M_W of each protein spot. By labeling every protein in this manner, data for the soldier caste could be compared to data for the worker caste. Manual comparison of spot data taken from the worker reference map (Fig. 7) yielded 577 differentially expressed proteins unique to workers while the soldier data from the soldier reference map (Fig. 37) yielded 517 differentially expressed proteins unique to soldiers. Table 6 includes the spot map number, experimental pI, experimental MW, protein identifier and caste for each differentially expressed protein yielded by these comparisons. Every effort was made to match spots between the caste reference maps to minimize duplication of spots in the tables. Many of the areas where differential expressed proteins were prevalent were inundated with an abundance of spots. As a result, the correct identity of the appropriate matched spot was obscured. To reduce the probability of mismatching spots between reference maps, questionable spots were treated as differentially expressed spots. Thus, it is likely some of the spots indicated as differentially expressed are actually matched spots between the reference maps. Further analysis using tighter pH ranges, additional mass spectrometry, or another method of protein identification would be required to confirm spots that should be matched.

Spots were selected from soldier gels and processed for mass spectrometry. Spots were taken from 3–10 pH range gels (Fig. 45) and 4–7 pH range gels (Fig. 46). Table 7 includes data for the putative identifications for some of the differentially expressed

proteins. Not all differentially expressed proteins were processed for mass spectrometry. The data set includes spot map number, experimental pI, experimental MW, protein identifier, and caste. The protein identifier was assigned using the following format; RF for *Reticulitermes flavipes*, S or W for soldier or worker caste, respectively, a four digit pI value, an underscore, and a six digit molecular weight value.

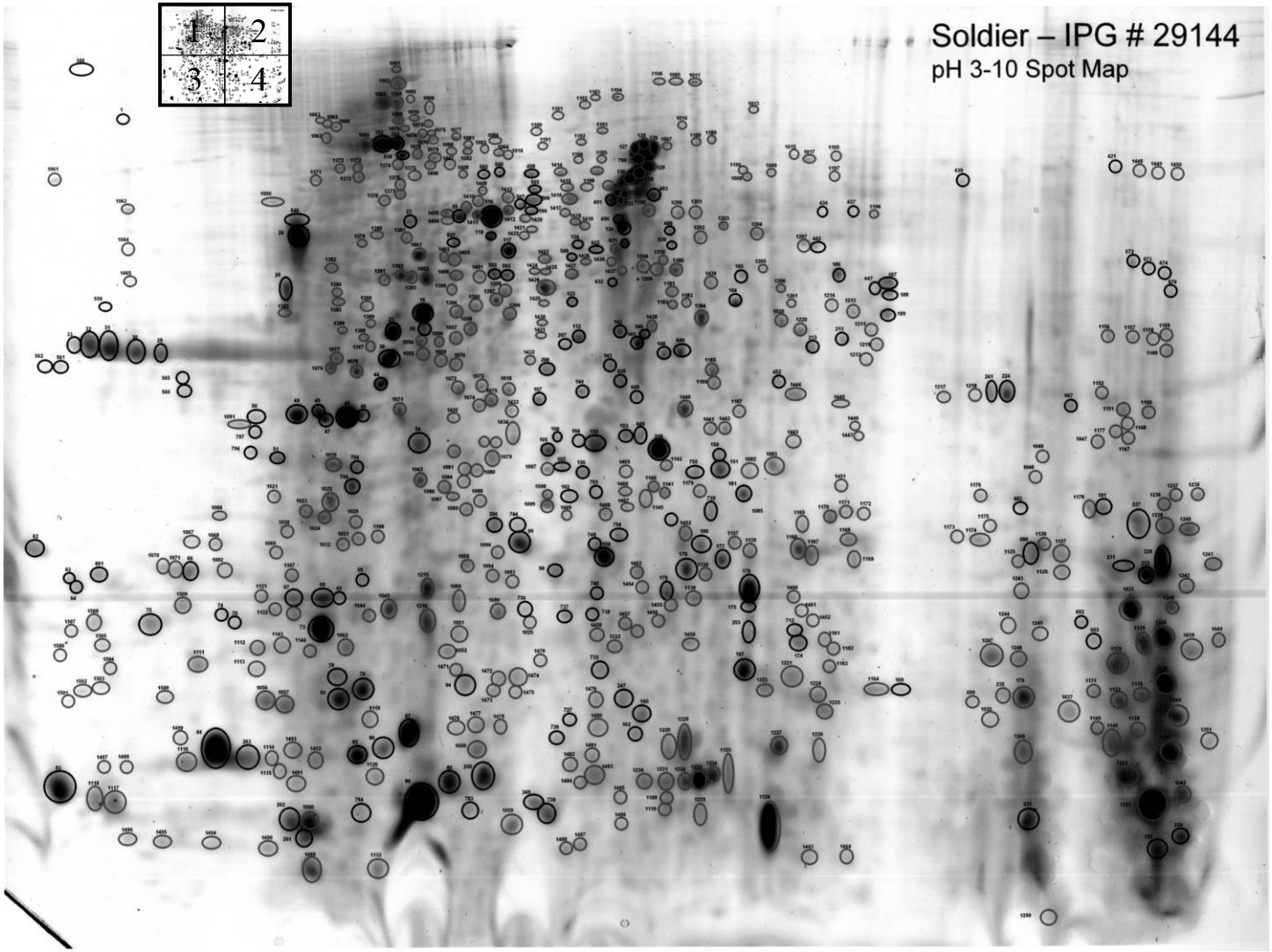


Fig. 37. Soldier caste reference map for *Reticulitermes flavipes*.
Inset: Overview of selected gel quadrants

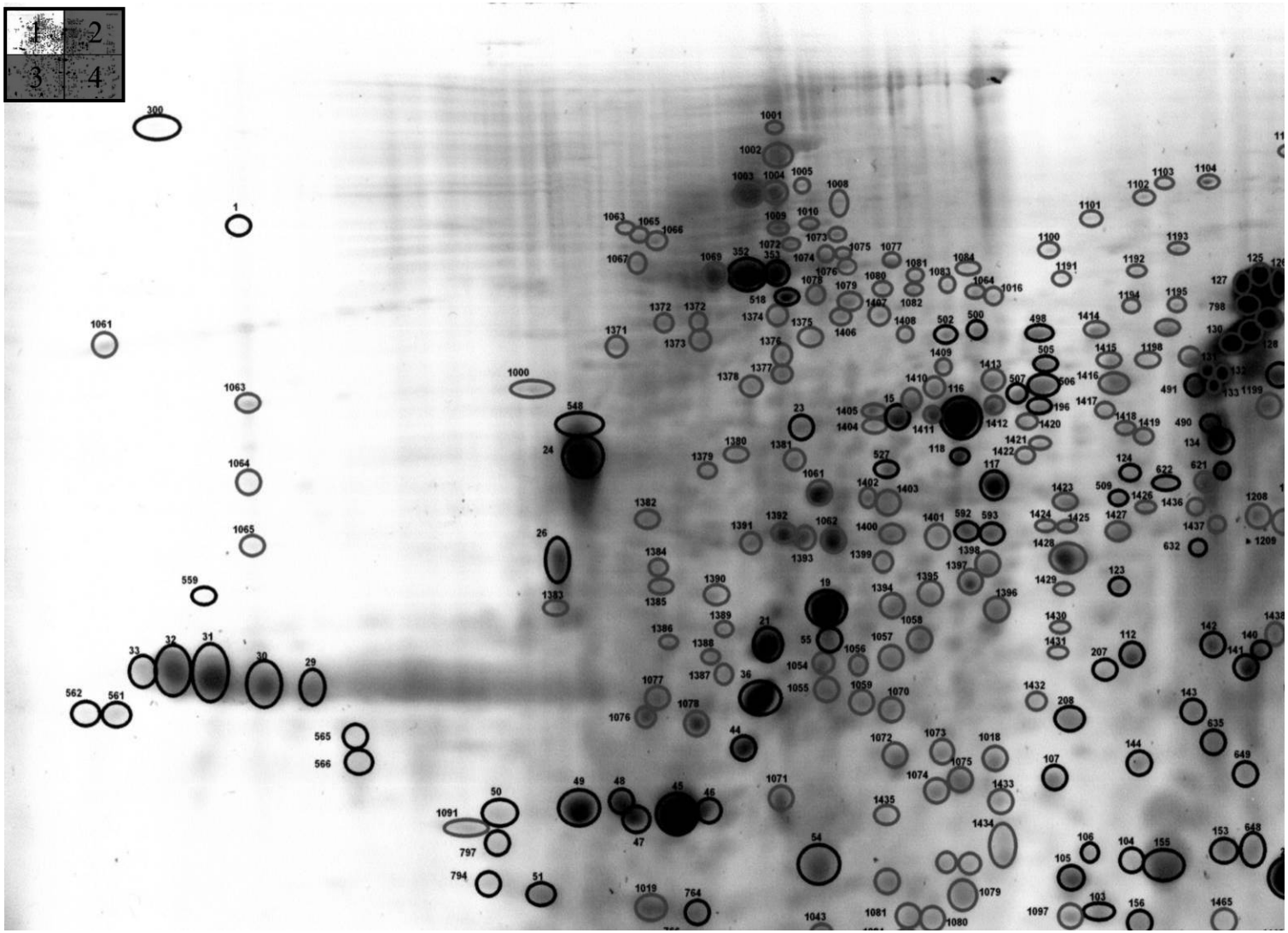


Fig. 38. Soldier caste reference map for *Reticulitermes flavipes* – Quadrant 1.
Inset: Highlight of selected gel quadrant.

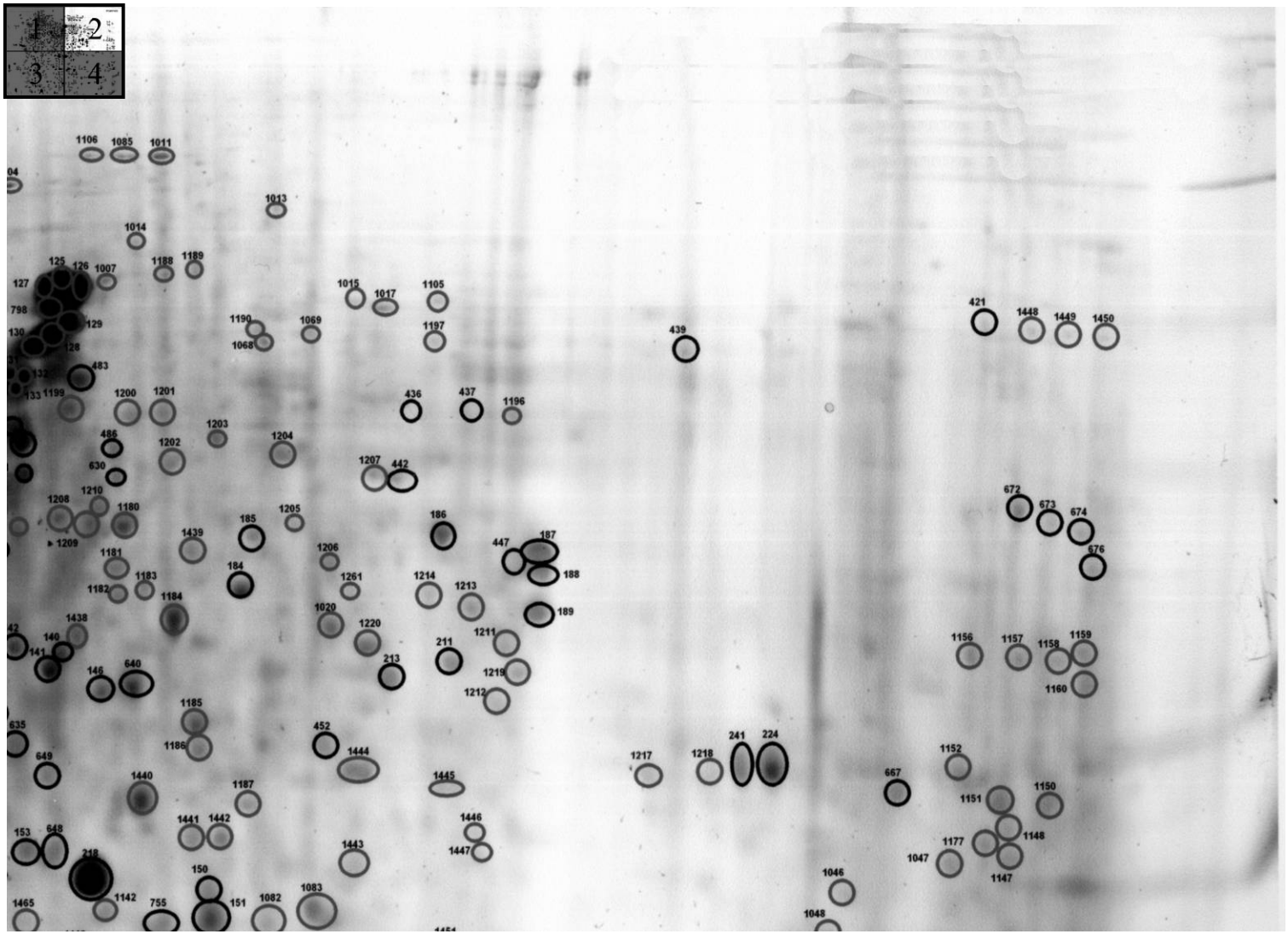


Fig. 39. Soldier caste reference map for *Reticulitermes flavipes* – Quadrant 2.
Inset: Highlight of selected gel quadrant.

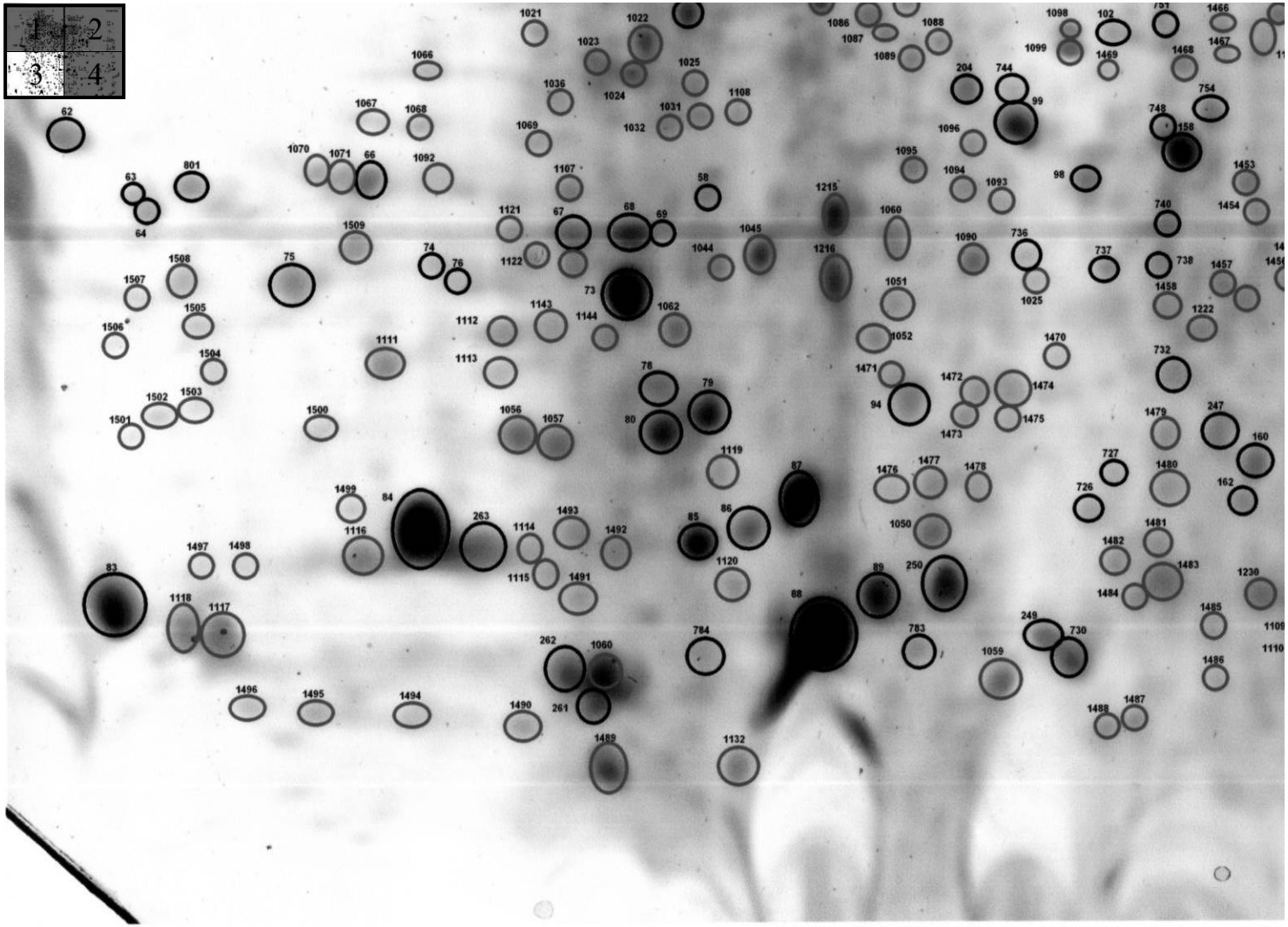


Fig. 40. Soldier caste reference map for *Reticulitermes flavipes* – Quadrant 3.
Inset: Highlight of selected gel quadrant.

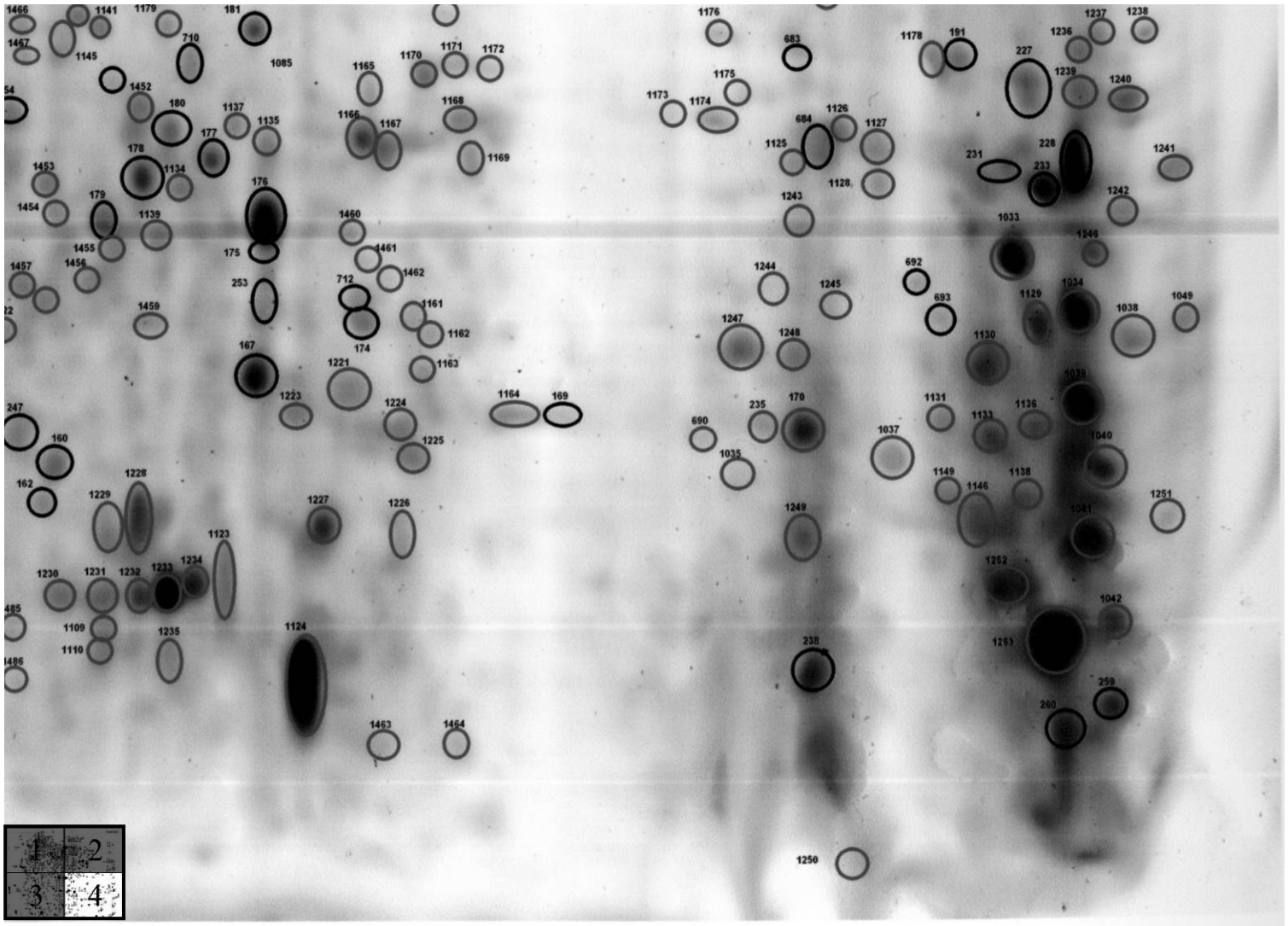


Fig. 41. Soldier caste reference map for *Reticulitermes flavipes* – Quadrant 4.
Inset: Highlight of selected gel quadrant.

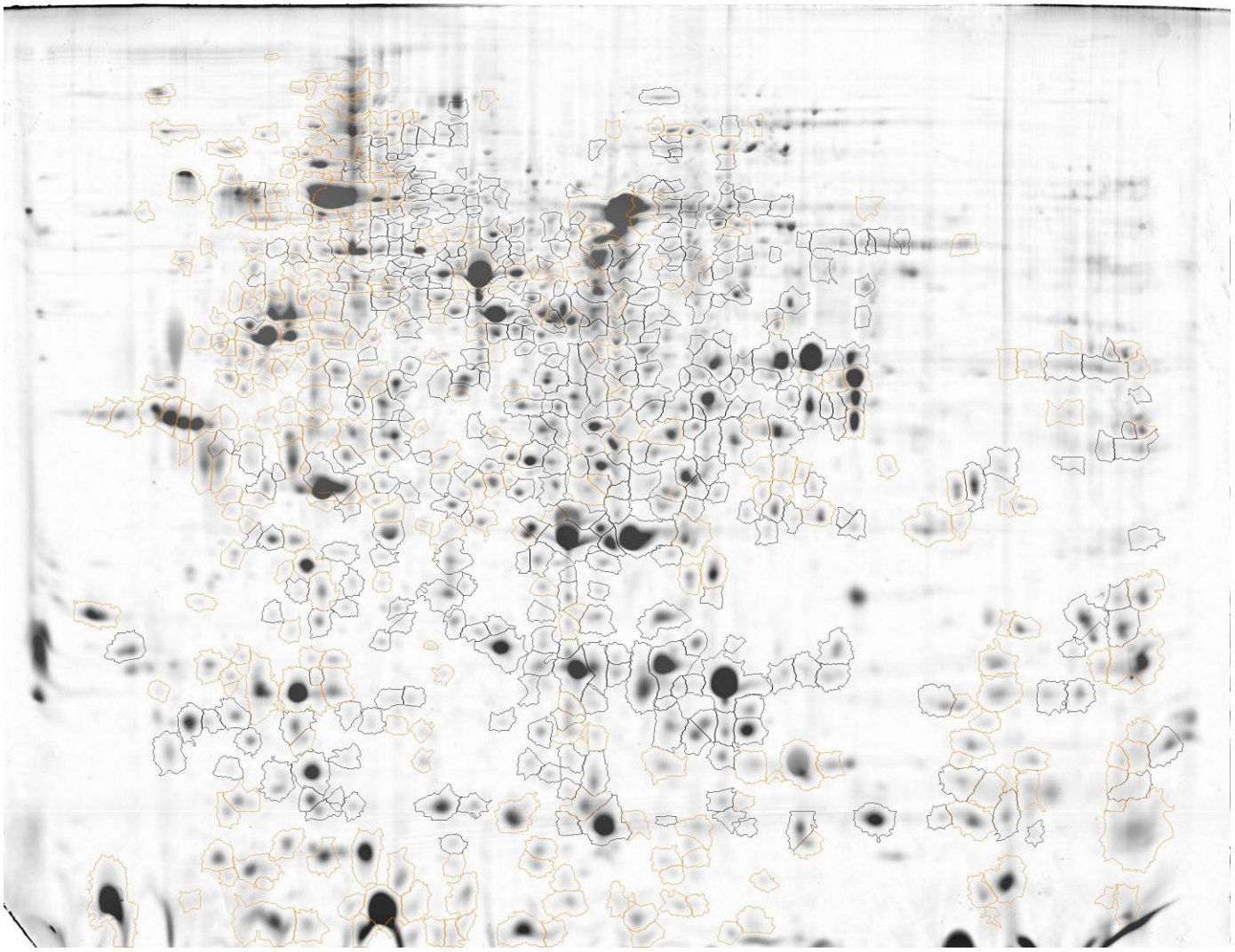


Fig. 42. Dymension image for common proteins between worker and soldier castes.

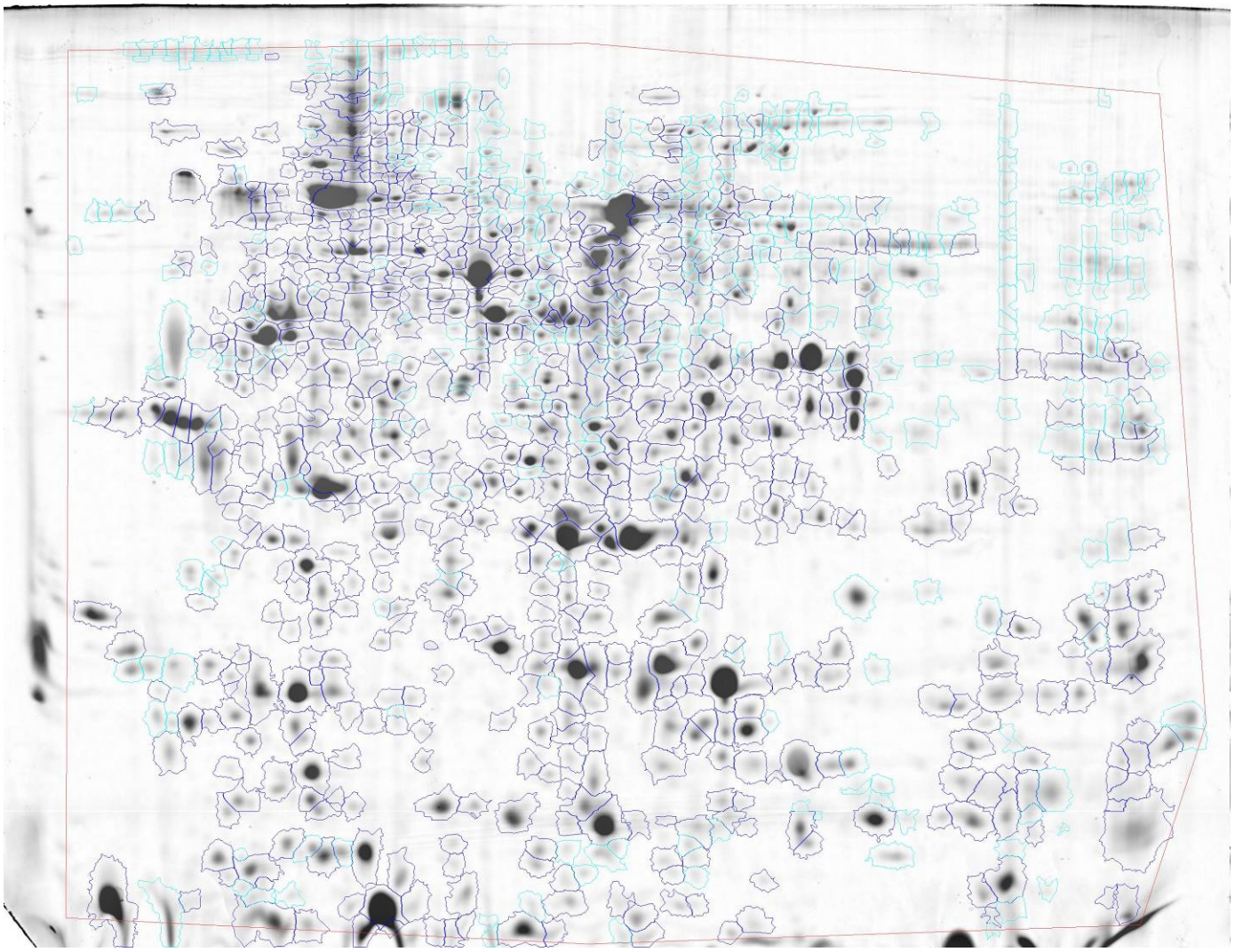


Fig. 43. Dymension image for ‘Worker Caste’.

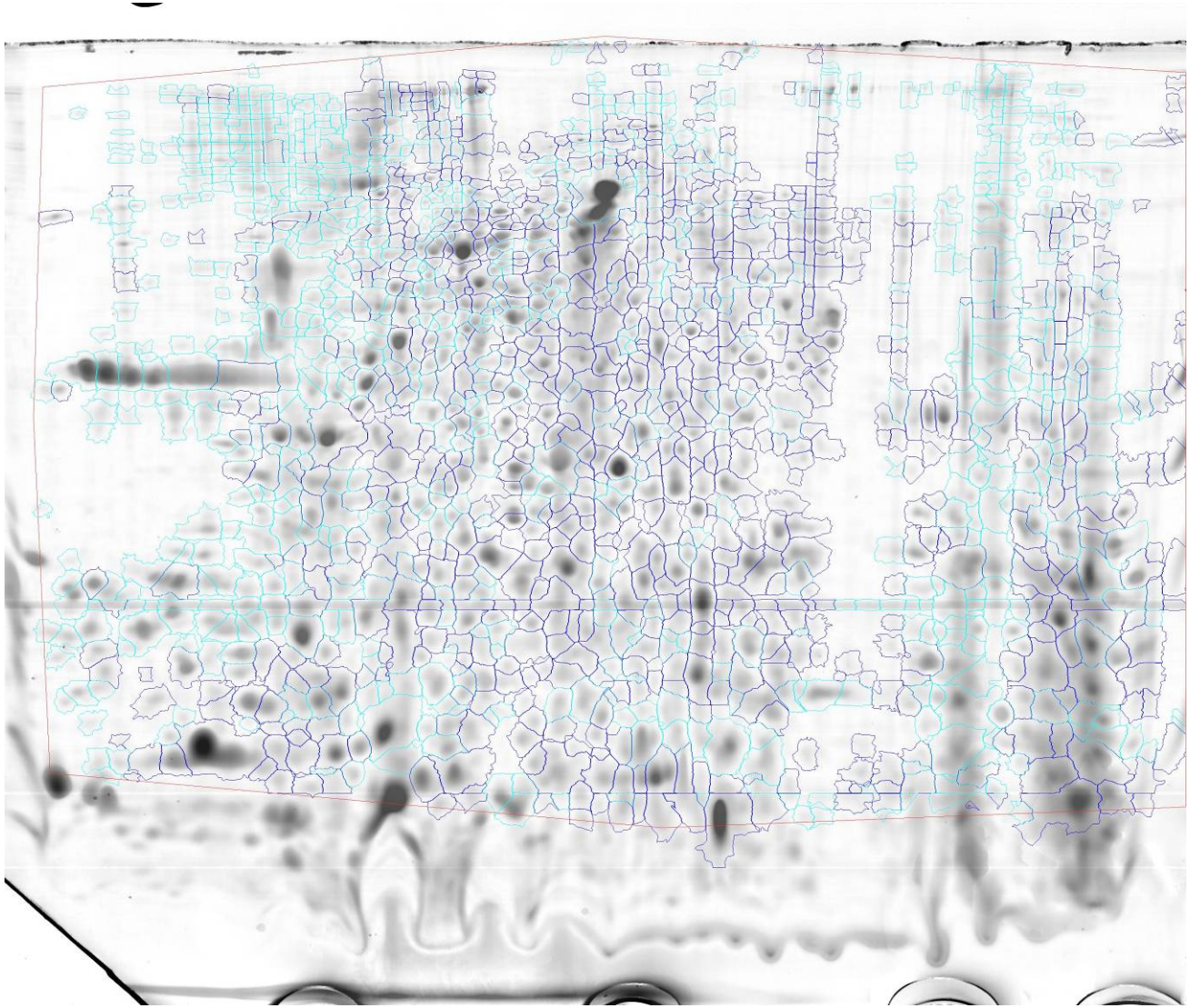


Fig. 44. Dymension image for 'Soldier Caste'.

Table 6. Differentially expressed proteins between worker and soldier castes from the same colony

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
2	3.94	89642	RFW0394_089642	Worker
3	3.77	72688	RFW0377_072688	Worker
4	3.49	67250	RFW0349_067250	Worker
5	4.01	67844	RFW0401_067844	Worker
6	4.72	100164	RFW0472_100164	Worker
7	4.54	97308	RFW0454_097308	Worker
8	4.44	81256	RFW0444_081256	Worker
9	4.72	79321	RFW0472_079321	Worker
10	4.86	77426	RFW0486_077426	Worker
11	4.59	53883	RFW0459_053883	Worker
12	4.63	54335	RFW0463_054335	Worker
13	4.61	56794	RFW0461_056794	Worker
14	4.81	64668	RFW0481_064668	Worker
16	5.35	57285	RFW0535_057285	Worker
17	4.87	47682	RFW0487_047682	Worker
18	5.19	41602	RFW0519_041602	Worker
20	5.20	34182	RFW0520_064182	Worker
22	4.99	35120	RFW0499_035120	Worker
25	4.48	39318	RFW0448_039318	Worker
27	4.69	39305	RFW0469_039305	Worker
28	4.46	33296	RFW0446_033296	Worker
34	3.78	35373	RFW0378_035373	Worker
35	4.91	28512	RFW0491_028512	Worker
37	5.14	33191	RFW0514_033191	Worker
38	5.28	27879	RFW0528_027879	Worker
39	5.35	29521	RFW0535_029521	Worker
40	5.62	26802	RFW0562_026802	Worker
41	5.48	25449	RFW0548_025449	Worker
42	5.43	24682	RFW0543_024682	Worker
43	5.31	24878	RFW0531_024878	Worker
52	4.68	20266	RFW0468_020266	Worker
53	5.04	22913	RFW0504_022913	Worker
56	5.14	17834	RFW0514_017834	Worker
57	5.43	17673	RFW0543_017676	Worker
59	4.74	18141	RFW0474_018141	Worker
60	4.03	18234	RFW0403_018234	Worker
61	4.00	16674	RFW0400_016674	Worker
65	3.64	14302	RFW0364_014302	Worker
70	4.84	13764	RFW0484_013764	Worker
71	4.88	15811	RFW0488_015811	Worker
72	4.95	14976	RFW0495_014976	Worker
77	4.22	13412	RFW0422_013412	Worker
81	4.29	12749	RFW0429_012749	Worker
82	4.20	12431	RFW0420_012431	Worker
90	5.52	15871	RFW0552_015871	Worker
91	5.57	14940	RFW0557_014940	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
92	5.73	16747	RFW0573_016747	Worker
93	5.71	13952	RFW0571_013952	Worker
95	5.92	8340	RFW0592_008340	Worker
96	5.92	8309	RFW0592_008309	Worker
97	6.29	13826	RFW0629_013826	Worker
100	5.73	20008	RFW0573_020008	Worker
101	5.82	20524	RFW0582_020524	Worker
108	5.81	25934	RFW0581_025934	Worker
109	5.81	26958	RFW0581_026958	Worker
110	5.91	25555	RFW0591_025555	Worker
111	5.81	25476	RFW0581_025476	Worker
113	5.95	36640	RFW0595_036640	Worker
114	5.54	35649	RFW0554_035649	Worker
115	5.61	39288	RFW0561_039288	Worker
119	5.08	49483	RFW0508_049483	Worker
120	5.24	44960	RFW0524_044960	Worker
121	6.09	50838	RFW0609_050838	Worker
122	5.71	33979	RFW0571_033979	Worker
135	5.57	38534	RFW0557_038534	Worker
136	6.01	53085	RFW0601_053085	Worker
137	5.75	36535	RFW0575_036535	Worker
138	5.85	35393	RFW0585_035393	Worker
139	5.98	33947	RFW0598_033947	Worker
145	6.13	26557	RFW0613_026557	Worker
147	6.40	29766	RFW0640_029766	Worker
148	6.49	29195	RFW0649_029195	Worker
149	6.40	28792	RFW0640_028792	Worker
152	6.20	25651	RFW0620_025651	Worker
154	6.12	25451	RFW0612_025451	Worker
157	6.40	18246	RFW0640_018246	Worker
159	6.64	14485	RFW0664_014485	Worker
161	6.51	7596	RFW0651_007596	Worker
163	6.90	6303	RFW0690_006303	Worker
164	6.87	6285	RFW0687_006285	Worker
165	6.84	7000	RFW0684_007000	Worker
166	6.84	8215	RFW0584_008215	Worker
168	7.60	8215	RFW0760_008215	Worker
169	8.04	8387	RFW0804_009387	Worker
169	8.04	8387	RFW0804_009387	Soldier
171	9.62	8008	RFW0962_008008	Worker
172	9.57	8214	RFW0957_008214	Worker
173	7.91	10224	RFW0791_010224	Worker
182	7.95	18524	RFW0795_018524	Worker
183	7.83	15508	RFW0783_015508	Worker
190	9.63	17770	RFW0963_017770	Worker
192	9.35	14594	RFW0935_014594	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
193	4.30	38056	RFW0430_038056	Worker
194	4.22	42676	RFW0422_042676	Worker
195	4.38	33683	RFW0438_033683	Worker
197	6.02	55139	RFW0602_055139	Worker
198	5.98	50376	RFW0598_050376	Worker
199	5.61	36563	RFW0561_036563	Worker
200	3.42	17442	RFW0342_017442	Worker
201	4.34	8866	RFW0434_008866	Worker
202	4.53	11862	RFW0453_011862	Worker
203	4.72	7518	RFW0472_007518	Worker
205	5.80	18695	RFW0580_018695	Worker
206	6.13	27540	RFW0613_027540	Worker
209	9.71	33764	RFW0971_033764	Worker
210	9.63	33718	RFW0963_033718	Worker
212	7.02	35105	RFW0702_035105	Worker
214	6.81	27015	RFW0681_027015	Worker
215	5.99	27069	RFW0599_027069	Worker
216	5.68	24570	RFW0568_024570	Worker
217	5.74	21583	RFW0574_021583	Worker
219	6.71	22934	RFW0671_002934	Worker
220	7.93	12907	RFW0793_012907	Worker
221	8.45	12627	RFW0845_012627	Worker
222	8.55	12625	RFW0855_012625	Worker
223	8.63	10918	RFW0863_010918	Worker
225	8.65	27104	RFW0865_027104	Worker
226	9.48	17693	RFW0948_017693	Worker
229	9.63	15030	RFW0963_015030	Worker
230	9.36	16435	RFW0936_016435	Worker
232	5.78	78451	RFW0578_078451	Worker
234	9.35	26747	RFW0935_026747	Worker
236	6.86	14473	RFW0686_014473	Worker
237	7.05	11794	RFW0705_011794	Worker
239	8.96	16832	RFW0896_016832	Worker
240	7.17	11658	RFW0717_011658	Worker
242	8.54	26664	RFW0854_026664	Worker
243	7.62	16178	RFW0762_016178	Worker
244	6.82	12745	RFW0682_012745	Worker
245	6.83	11854	RFW0683_011854	Worker
246	7.00	11230	RFW0700_011230	Worker
248	6.28	6097	RFW0628_006097	Worker
251	6.65	5927	RFW0665_006927	Worker
252	7.01	6153	RFW0701_006153	Worker
254	3.48	72819	RFW0348_072819	Worker
255	7.61	10115	RFW0761_010115	Worker
256	6.72	36296	RFW0672_036296	Worker
257	7.35	43835	RFW0735_043835	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
258	7.86	17151	RFW0786_017151	Worker
301	3.79	54512	RFW0379_054512	Worker
302	4.20	119431	RFW0420_119431	Worker
303	3.99	78190	RFW0399_078190	Worker
304	4.28	82721	RFW0428_082721	Worker
305	4.03	113256	RFW0403_113256	Worker
306	4.58	100162	RFW0458_100162	Worker
307	4.51	113508	RFW0451_113508	Worker
308	4.67	115078	RFW0467_115078	Worker
309	4.99	128812	RFW0499_128812	Worker
310	5.12	128112	RFW0512_128112	Worker
311	5.20	129680	RFW0520_129680	Worker
312	4.83	117520	RFW0483_117520	Worker
313	4.82	113811	RFW0482_113811	Worker
314	4.97	109150	RFW0497_109150	Worker
315	4.75	105532	RFW0475_105532	Worker
316	4.79	101123	RFW0479_101123	Worker
317	4.71	99483	RFW0471_099483	Worker
318	4.71	95444	RFW0471_095444	Worker
319	5.18	85288	RFW0518_085288	Worker
320	5.05	116786	RFW0505_116786	Worker
321	4.96	112305	RFW0496_112305	Worker
322	5.13	114446	RFW0513_114446	Worker
323	5.26	113362	RFW0526_113362	Worker
324	5.35	107518	RFW0535_107518	Worker
325	5.57	110319	RFW0557_110319	Worker
326	5.44	106726	RFW0544_106726	Worker
327	5.68	110332	RFW0568_110332	Worker
328	5.75	110038	RFW0575_110038	Worker
329	5.79	129057	RFW0579_129057	Worker
330	5.03	102473	RFW0503_102473	Worker
331	5.32	102920	RFW0532_102920	Worker
332	5.38	103049	RFW0538_103049	Worker
333	4.79	97162	RFW0479_097162	Worker
334	4.91	99907	RFW0491_099907	Worker
335	5.00	97274	RFW0500_097274	Worker
336	5.29	97073	RFW0529_097073	Worker
337	4.87	93650	RFW0487_093650	Worker
338	5.29	96664	RFW0529_096664	Worker
339	5.55	102108	RFW0555_102108	Worker
340	4.33	106304	RFW0433_106304	Worker
341	4.44	103375	RFW0444_103375	Worker
342	4.50	107821	RFW0450_107821	Worker
343	4.18	65459	RFW0418_065459	Worker
344	4.17	68702	RFW0417_068702	Worker
345	4.21	67227	RFW0421_067227	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
346	4.31	59040	RFW0431_059040	Worker
347	4.19	71363	RFW0419_071363	Worker
348	4.25	85253	RFW0425_085253	Worker
349	4.26	65970	RFW0426_065970	Worker
350	4.42	72147	RFW0431_061076	Worker
351	4.29	102142	RFW0429_102142	Worker
354	4.84	87102	RFW0484_087102	Worker
355	4.67	98010	RFW0467_098010	Worker
356	4.71	94688	RFW0471_094688	Worker
357	5.01	94368	RFW0501_094368	Worker
358	5.00	91403	RFW0500_091403	Worker
359	4.94	83341	RFW0494_083341	Worker
360	4.73	89320	RFW0473_089320	Worker
361	4.83	88024	RFW0483_088024	Worker
362	4.88	85175	RFW0488_085175	Worker
363	4.91	88198	RFW0491_088198	Worker
364	4.97	90900	RFW0497_090900	Worker
365	5.28	95136	RFW0528_095136	Worker
366	5.24	93051	RFW0524_093051	Worker
367	5.01	87439	RFW0501_087439	Worker
368	4.94	87439	RFW0494_087439	Worker
369	5.18	81724	RFW0518_081724	Worker
370	4.93	80596	RFW0493_080596	Worker
371	4.99	78613	RFW0499_078613	Worker
372	5.00	84221	RFW0500_084221	Worker
373	5.49	81753	RFW0549_081753	Worker
374	5.51	79210	RFW0551_079210	Worker
375	5.15	70034	RFW0515_070034	Worker
376	5.33	79767	RFW0533_079767	Worker
377	4.97	75314	RFW0497_075314	Worker
378	5.34	75036	RFW0534_075036	Worker
379	5.29	65037	RFW0529_065037	Worker
380	5.68	77245	RFW0568_077245	Worker
381	5.65	78977	RFW0565_078977	Worker
382	5.67	81583	RFW0567_081583	Worker
383	5.66	85064	RFW0566_085064	Worker
384	5.48	76176	RFW0548_076176	Worker
385	5.78	86944	RFW0578_086944	Worker
386	5.75	90695	RFW0575_090695	Worker
387	5.80	97153	RFW0580_097153	Worker
388	6.02	77018	RFW0602_077018	Worker
389	6.03	81456	RFW0603_081456	Worker
390	6.00	85010	RFW0600_085010	Worker
391	6.11	81611	RFW0611_081611	Worker
392	5.58	78807	RFW0558_078807	Worker
393	6.30	80854	RFW0630_080854	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
394	5.83	92622	RFW0583_092622	Worker
395	6.50	91768	RFW0650_091768	Worker
396	6.29	106872	RFW0629_106872	Worker
397	6.50	110202	RFW0650_110202	Worker
398	6.30	98774	RFW0630_098774	Worker
399	6.39	98539	RFW0639_098539	Worker
400	6.56	104127	RFW0656_104127	Worker
401	6.34	94909	RFW0634_094909	Worker
402	6.27	91622	RFW0627_091622	Worker
403	6.42	94031	RFW0642_094031	Worker
404	6.96	97563	RFW0696_97563	Worker
405	7.04	101143	RFW0704_101143	Worker
406	6.75	98219	RFW0675_098219	Worker
407	6.64	93090	RFW0664_093090	Worker
408	4.60	16618	RFW0460_016618	Worker
409	7.31	92018	RFW0731_092018	Worker
410	6.94	101280	RFW0694_101280	Worker
411	7.36	111390	RFW0736_111390	Worker
412	7.52	111390	RFW0752_111390	Worker
413	7.52	99323	RFW0752_099323	Worker
414	7.63	100454	RFW0763_100454	Worker
415	7.69	104761	RFW0769_104761	Worker
416	9.57	81072	RFW0957_081072	Worker
417	9.33	85928	RFW0933_085928	Worker
418	9.21	85802	RFW0921_085802	Worker
419	9.23	78136	RFW0923_078136	Worker
420	9.33	77272	RFW0933_077272	Worker
421	9.33	61680	RFW0933_061680	Worker
421	9.33	61680	RFW0933_061680	Soldier
422	8.59	65167	RFW0859_065167	Worker
423	8.36	65023	RFW0836_065023	Worker
424	8.27	65160	RFW0827_065160	Worker
425	8.02	65976	RFW0802_065976	Worker
426	7.84	66451	RFW0784_066451	Worker
427	7.97	66532	RFW0797_066532	Worker
428	7.94	69781	RFW0794_069781	Worker
429	7.78	74727	RFW0778_074727	Worker
430	6.95	74217	RFW0695_074217	Worker
431	7.56	69502	RFW0756_069502	Worker
432	6.88	71376	RFW0688_071376	Worker
433	7.35	70100	RFW0735_070100	Worker
434	7.39	66475	RFW0739_066475	Worker
435	7.43	67254	RFW0743_067254	Worker
438	7.25	60074	RFW0725_060074	Worker
440	7.58	56398	RFW0758_056398	Worker
441	8.06	54506	RFW0806_054506	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
443	7.49	49422	RFW0749_049422	Worker
444	7.51	51576	RFW0751_051576	Worker
445	8.04	47253	RFW0804_047253	Worker
446	4.50	17894	RFW0450_017894	Worker
448	7.18	33672	RFW0718_033672	Worker
449	6.91	34078	RFW0691_034078	Worker
450	6.66	35139	RFW0666_035139	Worker
451	6.78	32139	RFW0678_032139	Worker
453	6.54	33595	RFW0654_033595	Worker
454	6.46	34432	RFW0646_034432	Worker
455	6.47	35121	RFW0647_035121	Worker
456	6.53	36144	RFW0653_036144	Worker
457	6.30	37580	RFW0630_037580	Worker
458	7.04	48199	RFW0704_048199	Worker
459	6.36	55162	RFW0636_055162	Worker
460	7.29	55607	RFW0729_055607	Worker
461	7.18	56724	RFW0718_056724	Worker
462	6.13	60162	RFW0613_060162	Worker
463	6.55	62859	RFW0655_062859	Worker
464	6.10	57392	RFW0610_057392	Worker
465	6.00	67163	RFW0600_067163	Worker
466	6.95	72430	RFW0695_072430	Worker
467	7.01	72632	RFW0701_072632	Worker
468	6.32	69053	RFW0632_069053	Worker
469	6.50	70822	RFW0650_070822	Worker
470	6.63	77781	RFW0663_077781	Worker
471	6.45	74462	RFW0645_074462	Worker
472	7.13	72576	RFW0713_072576	Worker
473	7.07	74696	RFW0707_074696	Worker
474	6.12	75807	RFW0612_075807	Worker
475	7.11	66383	RFW0711_066383	Worker
476	6.36	65528	RFW0636_065528	Worker
477	7.17	59510	RFW0717_059510	Worker
478	6.78	72037	RFW0678_072037	Worker
479	5.84	68883	RFW0584_068883	Worker
480	6.85	69123	RFW0685_069123	Worker
481	5.87	63637	RFW0587_063637	Worker
482	5.43	63180	RFW0543_063180	Worker
484	5.44	40713	RFW0544_040713	Worker
485	5.48	39605	RFW0548_039605	Worker
487	5.69	61760	RFW0569_061760	Worker
488	5.78	58727	RFW0578_058727	Worker
489	6.04	58362	RFW0604_058362	Worker
492	5.45	67803	RFW0545_067803	Worker
493	6.35	69359	RFW0638_068796	Worker
494	6.37	72714	RFW0637_072714	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
495	5.48	71424	RFW0548_071424	Worker
496	5.39	49680	RFW0539_049680	Worker
497	6.01	70878	RFW0601_070878	Worker
499	5.84	69881	RFW0584_069881	Worker
501	5.67	74115	RFW0567_074115	Worker
503	5.10	66296	RFW0510_066296	Worker
504	4.91	69456	RFW0491_069456	Worker
508	5.30	45384	RFW0530_045384	Worker
510	5.26	47595	RFW0526_047595	Worker
511	5.30	59308	RFW0530_059308	Worker
512	5.26	60843	RFW0526_060843	Worker
513	4.63	58841	RFW0463_058841	Worker
514	4.52	62800	RFW0452_062800	Worker
515	4.61	75153	RFW0461_075153	Worker
516	4.91	72480	RFW0491_072480	Worker
517	4.53	81488	RFW0453_081488	Worker
519	4.69	74056	RFW0469_074056	Worker
520	4.83	76211	RFW0483_076211	Worker
521	4.83	74046	RFW0483_074046	Worker
522	5.10	69109	RFW0510_069109	Worker
523	4.75	70962	RFW0475_070962	Worker
524	4.70	66323	RFW0470_066323	Worker
526	5.02	63067	RFW0502_063067	Worker
528	5.11	45118	RFW0511_045118	Worker
529	5.04	44824	RFW0504_044824	Worker
530	5.05	43805	RFW0505_043805	Worker
531	4.98	44667	RFW0498_044667	Worker
532	5.01	43171	RFW0501_043171	Worker
533	4.88	45814	RFW0488_045814	Worker
534	4.83	46144	RFW0483_046144	Worker
535	4.74	48775	RFW0474_048775	Worker
536	4.74	58675	RFW0474_058675	Worker
537	4.60	50544	RFW0460_050544	Worker
538	4.54	53759	RFW0454_053759	Worker
539	4.56	50700	RFW0456_050700	Worker
540	4.42	51269	RFW0442_051269	Worker
541	4.12	50559	RFW0412_050559	Worker
542	4.20	46706	RFW0420_046706	Worker
543	4.10	57328	RFW0410_057328	Worker
544	4.42	47258	RFW0442_047258	Worker
545	4.67	54037	RFW0467_054037	Worker
546	4.74	55435	RFW0474_055435	Worker
547	4.85	44740	RFW0485_044740	Worker
548	4.55	43941	RFW0455_043941	Worker
548	4.55	43941	RFW0455_043941	Soldier
549	6.26	52838	RFW0626_052838	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
550	4.39	50866	RFW0439_050866	Worker
551	4.61	41950	RFW0461_041950	Worker
552	4.66	39914	RFW0466_039914	Worker
553	4.76	38096	RFW0476_038096	Worker
554	4.57	34854	RFW0457_034854	Worker
555	4.54	36759	RFW0454_036759	Worker
556	4.45	37117	RFW0445_037117	Worker
557	4.29	37078	RFW0429_037078	Worker
558	3.90	45318	RFW0390_045318	Worker
563	3.37	35760	RFW0337_035760	Worker
564	4.08	26329	RFW0408_026329	Worker
567	4.33	26057	RFW0433_026057	Worker
568	4.44	35099	RFW0444_035099	Worker
569	4.78	28485	RFW0478_027485	Worker
570	4.79	26268	RFW0479_026268	Worker
571	4.73	34666	RFW0473_034666	Worker
572	4.71	31765	RFW0471_031765	Worker
573	4.87	32871	RFW0487_032871	Worker
574	4.88	35826	RFW0488_035826	Worker
575	4.84	38094	RFW0484_038094	Worker
576	4.93	39530	RFW0493_039530	Worker
577	4.90	40875	RFW0490_040875	Worker
578	4.84	50895	RFW0484_050895	Worker
579	4.88	40411	RFW0488_040411	Worker
580	5.36	37441	RFW0536_037441	Worker
581	5.15	35492	RFW0515_035492	Worker
582	5.01	45637	RFW0501_045637	Worker
583	5.00	39455	RFW0500_039455	Worker
584	5.11	40286	RFW0511_040286	Worker
585	5.32	39909	RFW0532_039909	Worker
586	5.39	38305	RFW0539_039305	Worker
587	5.53	35296	RFW0553_035296	Worker
588	5.63	33455	RFW0563_033455	Worker
589	5.47	37803	RFW0547_037803	Worker
590	5.36	40050	RFW0536_040050	Worker
591	5.71	45957	RFW0571_045957	Worker
594	5.35	41323	RFW0535_041323	Worker
595	5.39	40616	RFW0539_040616	Worker
596	6.08	47398	RFW0608_047398	Worker
597	5.51	38213	RFW0551_038213	Worker
598	5.51	38146	RFW0551_038146	Worker
599	5.55	37258	RFW0555_037258	Worker
600	6.11	45044	RFW0611_045044	Worker
601	5.98	41532	RFW0598_041532	Worker
602	5.67	34392	RFW0567_034392	Worker
603	6.04	38249	RFW0604_038249	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
604	5.90	30820	RFW0590_030820	Worker
605	5.76	30070	RFW0576_030070	Worker
606	5.80	33214	RFW0580_033214	Worker
607	5.85	28469	RFW0585_028469	Worker
608	5.84	26704	RFW0584_026704	Worker
609	5.82	31568	RFW0582_31568	Worker
610	6.00	25980	RFW0600_025980	Worker
611	5.74	26127	RFW0574_026127	Worker
612	5.61	26776	RFW0561_026776	Worker
613	5.60	25415	RFW0560_025415	Worker
614	5.20	24598	RFW0520_024598	Worker
615	5.28	23822	RFW0528_023822	Worker
616	5.35	23520	RFW0535_023520	Worker
617	5.42	31827	RFW0542_031827	Worker
618	5.48	29617	RFW0548_029617	Worker
619	5.25	37925	RFW0525_037925	Worker
620	6.26	48285	RFW0626_048285	Worker
623	6.51	53188	RFW0651_053188	Worker
624	5.57	37567	RFW0557_037567	Worker
625	5.58	37431	RFW0558_037431	Worker
626	5.63	36457	RFW0563_036457	Worker
627	7.50	29173	RFW0750_029173	Worker
628	5.71	35080	RFW0571_035080	Worker
629	5.77	37101	RFW0577_037101	Worker
631	5.67	35503	RFW0567_035503	Worker
633	5.76	33415	RFW0576_033415	Worker
634	6.15	26122	RFW0615_026122	Worker
636	5.82	33034	RFW0582_033034	Worker
637	5.78	33745	RFW0578_033745	Worker
638	5.81	32868	RFW0581_032868	Worker
639	5.99	33082	RFW0599_033082	Worker
641	6.23	35396	RFW0623_035396	Worker
642	6.19	32963	RFW0619_032963	Worker
643	6.32	32089	RFW0632_032089	Worker
644	6.47	31525	RFW0647_031525	Worker
645	6.58	30076	RFW0658_030076	Worker
646	6.72	27842	RFW0672_027842	Worker
647	6.32	28359	RFW0632_028359	Worker
650	6.59	27964	RFW0659_027964	Worker
651	6.65	28945	RFW0665_028945	Worker
652	6.92	27980	RFW0692_027980	Worker
653	7.13	27555	RFW0713_027555	Worker
654	7.15	28309	RFW0715_028309	Worker
655	6.96	29337	RFW0696_029337	Worker
656	6.94	30298	RFW0694_030298	Worker
657	7.07	30212	RFW0707_030212	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
658	7.10	29382	RFW0710_029382	Worker
660	7.24	30077	RFW0724_030077	Worker
661	7.32	29321	RFW0732_029321	Worker
662	7.28	28023	RFW0728_028023	Worker
663	7.46	28054	RFW0746_028054	Worker
664	7.73	27639	RFW0773_027639	Worker
665	7.80	30625	RFW0780_030625	Worker
666	8.24	27795	RFW0824_027795	Worker
668	8.98	42723	RFW0898_042723	Worker
669	9.26	42417	RFW0926_042417	Worker
670	9.45	42067	RFW0945_042067	Worker
671	9.26	45305	RFW0926_045305	Worker
675	9.68	41883	RFW0968_041883	Worker
677	9.25	35633	RFW0925_035633	Worker
678	9.17	32321	RFW0917_032321	Worker
679	9.46	31977	RFW0946_031977	Worker
680	9.21	29537	RFW0921_029537	Worker
681	9.69	31468	RFW0969_031468	Worker
682	9.72	28152	RFW0972_028152	Worker
685	8.83	13297	RFW0883_013297	Worker
686	9.08	12757	RFW0908_012757	Worker
687	8.72	11672	RFW0872_011672	Worker
688	8.58	10334	RFW0858_010334	Worker
689	8.55	9712	RFW0855_009712	Worker
691	8.55	8549	RFW0855_008549	Worker
694	9.93	12102	RFW0993_012102	Worker
695	9.91	11307	RFW0991_011307	Worker
696	6.92	14190	RFW0692_014190	Worker
697	9.62	10530	RFW0962_010530	Worker
698	9.51	10401	RFW0951_010401	Worker
699	9.22	15253	RFW0922_015253	Worker
700	9.70	8398	RFW0970_008398	Worker
701	8.85	7094	RFW0885_007094	Worker
702	8.67	6500	RFW0867_006500	Worker
703	7.75	10146	RFW0775_010146	Worker
704	7.30	5910	RFW0730_005910	Worker
705	7.16	5632	RFW0716_005632	Worker
706	7.37	6807	RFW0737_006807	Worker
707	7.91	15697	RFW0791_015697	Worker
709	7.48	15454	RFW0748_015454	Worker
711	7.22	18276	RFW0722_018276	Worker
713	7.34	10187	RFW0734_010187	Worker
714	6.79	10191	RFW0679_010191	Worker
715	7.23	8998	RFW0723_008998	Worker
716	7.33	8263	RFW0733_008263	Worker
717	7.22	8463	RFW0722_008463	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
719	6.98	8049	RFW0698_008049	Worker
720	6.80	7374	RFW0680_007374	Worker
721	6.98	6901	RFW0698_006901	Worker
722	6.54	6207	RFW0654_006207	Worker
723	6.31	6914	RFW0631_006914	Worker
724	6.20	6696	RFW0620_006696	Worker
725	6.22	7582	RFW0622_007582	Worker
728	6.22	8189	RFW0622_008189	Worker
729	6.81	8439	RFW0681_008439	Worker
731	6.62	8207	RFW0662_008207	Worker
733	6.12	9777	RFW0612_009777	Worker
734	6.53	13843	RFW0653_013843	Worker
735	6.40	9853	RFW0640_009853	Worker
739	6.41	15533	RFW0641_015533	Worker
741	6.48	16416	RFW0648_016416	Worker
742	6.55	17772	RFW0655_017772	Worker
743	6.49	18790	RFW0649_018790	Worker
745	6.07	18581	RFW0607_018581	Worker
746	6.06	19817	RFW0606_019817	Worker
747	6.06	20696	RFW0606_020696	Worker
749	6.11	21036	RFW0611_021036	Worker
750	6.12	22236	RFW0612_022236	Worker
752	6.18	24126	RFW0618_024126	Worker
753	6.18	23085	RFW0618_023085	Worker
756	5.77	24275	RFW0577_024275	Worker
757	5.55	22975	RFW0555_022975	Worker
758	5.40	15529	RFW0540_015529	Worker
759	5.54	18480	RFW0554_018480	Worker
760	5.46	19120	RFW0546_019120	Worker
761	5.27	19057	RFW0527_019057	Worker
762	5.40	20023	RFW0540_020023	Worker
763	5.38	21154	RFW0538_021154	Worker
765	5.19	19722	RFW0519_019722	Worker
767	5.21	18948	RFW0521_018948	Worker
768	4.73	16893	RFW0473_016893	Worker
769	4.58	21138	RFW0458_021138	Worker
770	4.53	19851	RFW0453_019851	Worker
771	5.61	15683	RFW0561_015683	Worker
772	5.53	15295	RFW0553_015295	Worker
773	5.07	14985	RFW0507_014985	Worker
774	5.72	14811	RFW0572_014811	Worker
775	4.65	14199	RFW0465_014199	Worker
776	4.78	12436	RFW0478_012436	Worker
777	4.95	13653	RFW0495_013653	Worker
778	4.49	13182	RFW0449_013182	Worker
779	9.48	8823	RFW0948_008823	Worker

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
780	5.56	7703	RFW0556_007703	Worker
781	5.54	7023	RFW0554_007023	Worker
782	5.46	5618	RFW0516_005618	Worker
785	4.69	5748	RFW0469_005748	Worker
786	4.08	14331	RFW0408_014331	Worker
787	3.82	15487	RFW0382_015487	Worker
788	4.22	15729	RFW0422_015729	Worker
789	4.30	15106	RFW0430_015106	Worker
790	4.11	19281	RFW0411_019281	Worker
791	4.15	19892	RFW0415_019892	Worker
792	4.26	21523	RFW0426_021523	Worker
793	4.17	22523	RFW0417_022523	Worker
795	4.22	22326	RFW0422_022326	Worker
796	4.20	23609	RFW0420_023609	Worker
799	3.99	20451	RFW0399_020451	Worker
800	5.65	35360	RFW0565_035360	Worker
802	3.80	12075	RFW0380_012075	Worker
803	3.88	13882	RFW0388_013882	Worker
804	3.96	82270	RFW0396_082270	Worker
805	4.07	46526	RFW0407_046526	Worker
807	4.08	45002	RFW0408_045002	Worker
808	5.37	96442	RFW0537_096442	Worker
809	5.98	75042	RFW0598_075042	Worker
810	8.14	41720	RFW0814_041720	Worker
811	8.69	41062	RFW0869_041062	Worker
812	5.48	39057	RFW0548_039057	Worker
813	5.47	38612	RFW0547_038612	Worker
1000	4.35	58050	RFS0435_058050	Soldier
1001	4.94	91191	RFS0494_091191	Soldier
1002	4.98	84301	RFS0498_084301	Soldier
1003	4.91	79585	RFS0491_079585	Soldier
1004	5.00	77685	RFS0500_077685	Soldier
1005	5.06	80277	RFS0506_080277	Soldier
1006	4.83	64658	RFS0483_064658	Soldier
1007	6.72	76920	RFS0672_076920	Soldier
1008	5.31	77594	RFS0531_077594	Soldier
1009	5.02	73971	RFS0502_073971	Soldier
1010	5.12	75088	RFS0512_075088	Soldier
1011	6.83	94615	RFS0683_094615	Soldier
1012	6.34	64815	RFS0634_064815	Soldier
1013	7.15	83862	RFS0715_083862	Soldier
1014	6.78	81850	RFS0678_081850	Soldier
1015	7.44	71500	RFS0744_071500	Soldier
1017	7.54	70340	RFS0754_070340	Soldier
1018	5.83	31338	RFS0583_034338	Soldier
1019	4.80	22188	RFS0480_022188	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1020	7.33	36643	RFS0733_036643	Soldier
1021	4.53	18532	RFS0453_018532	Soldier
1022	4.78	19577	RFS0478_019577	Soldier
1023	4.68	17647	RFS0468_017647	Soldier
1024	4.79	17718	RFS0479_017718	Soldier
1025	4.95	17963	RFS0495_017963	Soldier
1026	5.82	11559	RFS0582_011559	Soldier
1027	6.72	15849	RFS0672_015849	Soldier
1028	4.50	6075	RFS0450_006075	Soldier
1029	4.62	5952	RFS0462_005952	Soldier
1030	5.69	4779	RFS0569_004779	Soldier
1031	4.99	16538	RFS0499_016538	Soldier
1032	4.88	15938	RFS0488_015938	Soldier
1033	9.41	16088	RFS0941_016088	Soldier
1034	9.63	14651	RFS0963_014651	Soldier
1035	8.55	9985	RFS0855_009985	Soldier
1036	4.53	15977	RFS0453_015977	Soldier
1037	9.02	10741	RFS0902_010741	Soldier
1038	9.81	13974	RFS0981_013974	Soldier
1039	9.59	12237	RFS0959_012237	Soldier
1040	6.94	10567	RFS0694_010567	Soldier
1041	9.62	9970	RFS0962_009970	Soldier
1042	9.76	8064	RFS0976_008064	Soldier
1043	5.27	21868	RFS0527_021868	Soldier
1044	5.05	10928	RFS0505_010928	Soldier
1045	5.20	12379	RFS0520_012379	Soldier
1046	8.91	25907	RFS0891_025907	Soldier
1047	9.19	28352	RFS0919_028352	Soldier
1048	8.81	23617	RFS0881_023617	Soldier
1049	9.98	14676	RFS0998_14676	Soldier
1050	5.54	6007	RFS0554_006007	Soldier
1051	5.56	11507	RFS0556_011507	Soldier
1052	5.49	10356	RFS0549_010356	Soldier
1054	5.24	35599	RFS0524_035599	Soldier
1055	5.31	33620	RFS0531_033620	Soldier
1056	4.50	6075	RFS0450_006075	Soldier
1057	4.62	5952	RFS0462_005952	Soldier
1058	5.46	36198	RFS0546_036198	Soldier
1059	5.92	5631	RFS0592_005631	Soldier
1060	5.53	12716	RFS0553_012716	Soldier
1061	5.24	49485	RFS0524_049485	Soldier
1062	5.28	45227	RFS0528_045227	Soldier
1063	4.56	73521	RFS0456_073521	Soldier
1064	3.53	48523	RFS0353_048523	Soldier
1065	4.64	72762	RFS0464_072762	Soldier
1066	4.70	72373	RFS0470_072373	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1067	4.66	70089	RFS0466_070089	Soldier
1068	7.16	66996	RFS0716_066996	Soldier
1069	7.30	68467	RFS0730_068467	Soldier
1070	5.49	33225	RFS0549_033225	Soldier
1071	5.12	26963	RFS0512_026963	Soldier
1072	5.49	30055	RFS0549_030055	Soldier
1073	5.67	31037	RFS0567_031037	Soldier
1074	5.59	28103	RFS0559_028103	Soldier
1075	5.73	29319	RFS0573_029319	Soldier
1076	4.71	32377	RFS0471_032377	Soldier
1077	4.74	34154	RFS0474_034154	Soldier
1078	4.83	32546	RFS0483_032546	Soldier
1079	5.70	23769	RFS0570_023769	Soldier
1080	5.60	23097	RFS0560_023097	Soldier
1081	5.53	22787	RFS0553_022787	Soldier
1082	7.16	19371	RFS0716_019371	Soldier
1083	7.35	18646	RFS0735_018646	Soldier
1084	5.53	22787	RFS0553_022787	Soldier
1085	6.53	11327	RFS0653_011327	Soldier
1086	5.40	21588	RFS0540_021588	Soldier
1087	5.46	20700	RFS0546_020700	Soldier
1088	5.61	20207	RFS0561_020207	Soldier
1089	5.54	19408	RFS0554_019408	Soldier
1090	6.53	11327	RFS0653_011327	Soldier
1091	4.33	25338	RFS0433_025338	Soldier
1092	4.10	9786	RFS0410_009786	Soldier
1093	5.81	12997	RFS0581_012997	Soldier
1094	5.72	13546	RFS0572_013546	Soldier
1095	5.57	14912	RFS0557_014912	Soldier
1096	5.71	14614	RFS0571_014614	Soldier
1097	5.99	22355	RFS0599_022355	Soldier
1098	5.97	20186	RFS0597_020186	Soldier
1099	5.99	18445	RFS0599_018445	Soldier
1100	5.88	75839	RFS0588_075839	Soldier
1101	6.00	80283	RFS0600_080283	Soldier
1102	6.08	83773	RFS0608_083773	Soldier
1103	6.20	85822	RFS0620_085822	Soldier
1104	6.32	87145	RFS0632_087145	Soldier
1105	7.69	71402	RFS0769_071402	Soldier
1106	6.57	92807	RFS0657_092807	Soldier
1107	4.62	12537	RFS0462_012537	Soldier
1108	5.08	17087	RFS0508_017087	Soldier
1109	6.65	7556	RFS0665_007556	Soldier
1110	6.65	7222	RFS0665_007222	Soldier
1111	4.16	7326	RFS0416_007326	Soldier
1112	4.47	7672	RFS0447_007672	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1113	4.47	6962	RFS0447_006962	Soldier
1114	4.55	4673	RFS0455_004673	Soldier
1115	4.63	4437	RFS0463_004437	Soldier
1116	4.06	5147	RFS0406_005147	Soldier
1117	3.75	5631	RFS0375_005631	Soldier
1118	3.65	5426	RFS0365_005426	Soldier
1119	4.86	6999	RFS0486_006999	Soldier
1120	5.03	4677	RFS0503_004677	Soldier
1121	4.49	10760	RFS0449_010760	Soldier
1122	4.55	9617	RFS0455_009617	Soldier
1123	7.06	8391	RFS0706_008391	Soldier
1124	7.30	7325	RFS0730_007325	Soldier
1125	8.70	17844	RFS0870_017844	Soldier
1126	8.86	19340	RFS0886_019340	Soldier
1127	8.98	19348	RFS0898_019348	Soldier
1128	8.99	17713	RFS0899_017713	Soldier
1129	9.44	14261	RFS0944_014261	Soldier
1130	9.42	11531	RFS0942_011531	Soldier
1131	9.15	11393	RFS0915_011393	Soldier
1132	5.01	3307	RFS0501_003307	Soldier
1133	9.30	11080	RFS0930_011080	Soldier
1134	6.96	13307	RFS0696_013307	Soldier
1135	7.17	13639	RFS0717_013639	Soldier
1136	9.42	11531	RFS0942_011531	Soldier
1137	7.08	14569	RFS0708_014569	Soldier
1138	9.36	9639	RFS0936_009639	Soldier
1139	6.86	12551	RFS0686_012551	Soldier
1140	6.62	18836	RFS0662_018836	Soldier
1141	6.69	18031	RFS0669_018031	Soldier
1142	6.68	20188	RFS0668_020188	Soldier
1143	4.60	7609	RFS0460_007609	Soldier
1144	4.75	7711	RFS0475_007711	Soldier
1145	6.57	17044	RFS0657_017044	Soldier
1146	9.24	9572	RFS0924_009572	Soldier
1147	9.39	27949	RFS0939_027949	Soldier
1148	9.41	28922	RFS0941_028922	Soldier
1149	9.15	10109	RFS0915_010109	Soldier
1150	9.51	32211	RFS0951_032211	Soldier
1151	9.36	32293	RFS0936_032293	Soldier
1152	9.18	32621	RFS0918_032621	Soldier
1153	5.53	13266	RFS0553_013266	Soldier
1154	5.28	45227	RFS0528_045227	Soldier
1155	5.61	70264	RFS0561_070264	Soldier
1156	9.26	41425	RFS0926_041425	Soldier
1157	9.39	41205	RFS0939_041205	Soldier
1158	9.53	40917	RFS0953_040917	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1159	9.63	42211	RFS0963_042211	Soldier
1160	9.64	38861	RFS0964_038861	Soldier
1161	7.58	10775	RFS0758_010775	Soldier
1162	7.63	10537	RFS0763_010537	Soldier
1163	7.63	10128	RFS0763_010128	Soldier
1164	7.88	9821	RFS0788_009821	Soldier
1165	7.47	14465	RFS0747_014465	Soldier
1166	7.44	13311	RFS0744_013311	Soldier
1167	7.54	12960	RFS0754_012960	Soldier
1168	7.73	13401	RFS0773_013401	Soldier
1169	7.76	12637	RFS0776_012637	Soldier
1170	7.64	14404	RFS0764_014404	Soldier
1171	7.75	14811	RFS0775_014811	Soldier
1172	7.86	14894	RFS0786_014894	Soldier
1173	8.36	16864	RFS0836_016864	Soldier
1174	8.35	17974	RFS0835_017974	Soldier
1175	8.59	19251	RFS0859_017251	Soldier
1176	8.56	20754	RFS0856_020754	Soldier
1177	9.31	29145	RFS0931_029145	Soldier
1178	9.17	21553	RFS0917_021553	Soldier
1179	6.86	17869	RFS0686_017869	Soldier
1180	6.70	47754	RFS0575_036535	Soldier
1181	6.67	43864	RFS0667_043864	Soldier
1182	6.68	41592	RFS0668_041592	Soldier
1183	6.75	42007	RFS0675_042007	Soldier
1184	6.86	39209	RFS0686_039209	Soldier
1185	6.93	31940	RFS0693_031940	Soldier
1186	6.93	29506	RFS0693_029506	Soldier
1187	7.10	25422	RFS0710_025422	Soldier
1188	6.87	76462	RFS0687_076462	Soldier
1189	6.95	78319	RFS0695_078319	Soldier
1190	7.13	69123	RFS0713_069123	Soldier
1191	5.94	73115	RFS0594_073115	Soldier
1192	6.15	74789	RFS0615_074789	Soldier
1193	6.26	76985	RFS0626_076985	Soldier
1194	6.14	71328	RFS0614_071328	Soldier
1195	6.26	72119	RFS0626_072119	Soldier
1196	7.87	58045	RFS0787_058045	Soldier
1197	7.67	66218	RFS0767_066218	Soldier
1198	6.20	64742	RFS0620_064742	Soldier
1199	6.60	62070	RFS0660_062070	Soldier
1200	6.74	58227	RFS0674_058227	Soldier
1201	6.83	58142	RFS0683_057142	Soldier
1202	6.83	52172	RFS0683_052172	Soldier
1203	6.96	56847	RFS0696_056847	Soldier
1204	7.26	53114	RFS0726_053114	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1205	7.29	45081	RFS0729_045081	Soldier
1206	7.33	42339	RFS0733_042339	Soldier
1207	7.47	51721	RFS0747_051721	Soldier
1208	6.50	47496	RFS0650_047496	Soldier
1209	6.60	46963	RFS0660_046963	Soldier
1210	6.63	49616	RFS0663_049616	Soldier
1211	7.86	35631	RFS0786_035631	Soldier
1212	7.85	30846	RFS0785_030846	Soldier
1213	7.75	37492	RFS0775_037492	Soldier
1214	7.63	38018	RFS0763_038018	Soldier
1215	5.37	14014	RFS0537_014014	Soldier
1216	5.38	11466	RFS0538_011466	Soldier
1217	8.31	28702	RFS0831_028702	Soldier
1218	8.48	60279	RFS0848_060279	Soldier
1219	7.89	33879	RFS0789_033879	Soldier
1220	7.45	35273	RFS0745_035273	Soldier
1221	7.41	10268	RFS0741_010268	Soldier
1222	6.40	10859	RFS0640_010859	Soldier
1223	7.27	10173	RFS0727_010173	Soldier
1224	7.51	9681	RFS0751_009681	Soldier
1225	7.60	9355	RFS0760_009355	Soldier
1226	7.55	8460	RFS0755_009460	Soldier
1227	7.31	8888	RFS0731_008888	Soldier
1228	6.87	9237	RFS0687_009237	Soldier
1229	6.79	9046	RFS0679_009046	Soldier
1230	6.55	8178	RFS0655_008178	Soldier
1231	6.68	8182	RFS0668_008182	Soldier
1232	6.79	8275	RFS0679_008275	Soldier
1233	6.88	8343	RFS0688_008343	Soldier
1234	6.97	8391	RFS0697_008391	Soldier
1235	6.91	7441	RFS0691_007441	Soldier
1236	9.64	24058	RFS0964_024058	Soldier
1237	9.66	25433	RFS0966_025433	Soldier
1238	9.83	25205	RFS0983_025205	Soldier
1239	9.62	22168	RFS0962_022168	Soldier
1240	9.80	22017	RFS0980_022017	Soldier
1241	9.94	19832	RFS0994_019832	Soldier
1242	9.83	17565	RFS0983_017565	Soldier
1243	8.73	15834	RFS0873_015834	Soldier
1244	8.53	12990	RFS0853_012990	Soldier
1245	8.81	13505	RFS0881_013505	Soldier
1246	9.66	16146	RFS0966_016146	Soldier
1247	8.53	11966	RFS0853_011966	Soldier
1248	8.73	12355	RFS0873_012355	Soldier
1249	8.74	9352	RFS0874_009352	Soldier
1251	9.80	9906	RFS0980_009906	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1252	9.35	8636	RFS0935_008636	Soldier
1253	9.52	7823	RFS0952_007823	Soldier
1254	6.37	50573	RFS0637_050573	Soldier
1255	4.73	62749	RFS0473_062749	Soldier
1256	4.68	54868	RFS0468_054868	Soldier
1257	4.69	57069	RFS0469_057069	Soldier
1258	4.68	54868	RFS0468_054868	Soldier
1259	4.70	29745	RFS0470_029745	Soldier
1260	5.31	49981	RFS0531_049981	Soldier
1261	7.28	45603	RFS0728_045603	Soldier
1262	3.90	12734	RFS0390_012734	Soldier
1263	4.90	78351	RFS0490_078351	Soldier
1264	4.97	94994	RFS0497_094994	Soldier
1265	4.92	79680	RFS0492_079680	Soldier
1266	5.02	81739	RFS0502_081739	Soldier
1267	5.07	79888	RFS0507_079888	Soldier
1269	5.06	85569	RFS0506_085569	Soldier
1270	5.19	75459	RFS0519_075459	Soldier
1271	5.02	75350	RFS0502_075350	Soldier
1272	4.94	65304	RFS0494_065304	Soldier
1273	4.92	79680	RFS0492_079680	Soldier
1274	4.82	45011	RFS0482_045011	Soldier
1275	4.74	41957	RFS0474_041957	Soldier
1276	4.80	41907	RFS0480_041907	Soldier
1277	4.87	41942	RFS0487_041942	Soldier
1278	4.87	37818	RFS0487_037818	Soldier
1279	5.00	49040	RFS0500_049040	Soldier
1281	5.14	59222	RFS0514_059222	Soldier
1282	5.11	52891	RFS0511_052891	Soldier
1283	5.18	41114	RFS0518_041114	Soldier
1284	5.14	39563	RFS0514_039563	Soldier
1285	5.16	37388	RFS0516_037388	Soldier
1287	5.51	45963	RFS0551_045963	Soldier
1288	5.50	46713	RFS0550_046713	Soldier
1289	5.44	30489	RFS0544_030489	Soldier
1290	5.34	29758	RFS0534_029758	Soldier
1291	5.28	29129	RFS0528_029129	Soldier
1292	5.17	26047	RFS0517_026047	Soldier
1293	5.58	23913	RFS0558_023913	Soldier
1293	6.76	6093	RFS0676_006093	Soldier
1294	6.19	9534	RFS0619_009534	Soldier
1295	6.30	9762	RFS0630_009762	Soldier
1297	5.85	5413	RFS0585_005413	Soldier
1298	5.93	6975	RFS0593_006975	Soldier
1299	6.01	6851	RFS0601_006851	Soldier
1300	5.93	5121	RFS0593_005121	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1301	5.75	9762	RFS0575_009762	Soldier
1302	5.35	15710	RFS0535_015710	Soldier
1303	5.36	13312	RFS0536_013312	Soldier
1304	5.14	14793	RFS0514_014793	Soldier
1305	5.53	9982	RFS0553_009982	Soldier
1306	5.31	8070	RFS0531_008070	Soldier
1307	5.24	8222	RFS0524_008222	Soldier
1308	5.17	8364	RFS0517_008364	Soldier
1309	5.30	5592	RFS0530_005592	Soldier
1310	5.17	5578	RFS0517_005578	Soldier
1311	5.16	6301	RFS0516_006301	Soldier
1312	4.77	9769	RFS0477_009769	Soldier
1313	4.94	14333	RFS0494_014333	Soldier
1314	4.74	13222	RFS0474_013222	Soldier
1315	4.76	11756	RFS0476_011756	Soldier
1316	4.59	9854	RFS0459_009854	Soldier
1317	5.24	5874	RFS0524_005874	Soldier
1318	4.68	5843	RFS0468_005843	Soldier
1319	4.74	5805	RFS0474_005805	Soldier
1320	4.69	4986	RFS0469_004986	Soldier
1321	4.23	12311	RFS0423_012311	Soldier
1322	4.13	12564	RFS0413_012564	Soldier
1323	4.39	14907	RFS0439_014907	Soldier
1324	4.36	16062	RFS0436_016062	Soldier
1325	4.16	15149	RFS0416_015149	Soldier
1326	4.00	17141	RFS0400_017141	Soldier
1327	3.87	15354	RFS0387_015354	Soldier
1328	4.18	17213	RFS0418_017213	Soldier
1329	4.10	14939	RFS0410_014939	Soldier
1330	4.11	13436	RFS0411_013436	Soldier
1331	4.87	5724	RFS0487_005724	Soldier
1332	4.57	50612	RFS0457_050612	Soldier
1333	4.97	46237	RFS0497_046237	Soldier
1334	4.56	24137	RFS0456_024137	Soldier
1335	4.68	23173	RFS0468_023173	Soldier
1336	6.16	15009	RFS0616_015009	Soldier
1337	6.05	18579	RFS0605_018579	Soldier
1338	5.28	29506	RFS0528_029506	Soldier
1339	5.63	29160	RFS0563_029160	Soldier
1340	6.44	18343	RFS0644_018343	Soldier
1341	6.77	22953	RFS0677_022953	Soldier
1342	6.77	25638	RFS0677_025638	Soldier
1343	6.84	33522	RFS0684_033522	Soldier
1344	6.25	28321	RFS0625_028321	Soldier
1345	6.54	52642	RFS0654_052642	Soldier
1346	6.54	47834	RFS0654_047834	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1347	6.45	59428	RFS0645_059428	Soldier
1348	6.41	58587	RFS0641_058587	Soldier
1349	6.36	58473	RFS0636_058473	Soldier
1350	6.35	61006	RFS0635_061006	Soldier
1351	5.16	61234	RFS0516_061234	Soldier
1352	6.09	64784	RFS0609_064784	Soldier
1353	6.60	6636	RFS0660_006636	Soldier
1354	6.17	54588	RFS0617_054588	Soldier
1355	6.22	52820	RFS0622_052820	Soldier
1356	6.28	49197	RFS0628_049197	Soldier
1357	5.76	42231	RFS0576_042231	Soldier
1358	5.67	43247	RFS0567_043247	Soldier
1359	5.69	39750	RFS0569_039750	Soldier
1360	5.70	36013	RFS0570_036013	Soldier
1371	4.58	62489	RFS0458_062489	Soldier
1372	4.68	64383	RFS0468_064383	Soldier
1373	4.86	62634	RFS0486_062634	Soldier
1374	4.99	63891	RFS0499_063891	Soldier
1375	5.20	63610	RFS0520_063610	Soldier
1376	5.09	63359	RFS0509_063359	Soldier
1377	5.10	59957	RFS0510_059957	Soldier
1378	5.00	58650	RFS0500_059650	Soldier
1379	4.97	52377	RFS0497_052377	Soldier
1380	5.06	52975	RFS0506_052975	Soldier
1381	5.14	51979	RFS0514_051979	Soldier
1382	4.74	46353	RFS0474_046353	Soldier
1383	4.51	38759	RFS0451_038759	Soldier
1384	4.76	42683	RFS0476_042683	Soldier
1385	4.76	41191	RFS0476_041191	Soldier
1386	4.79	36721	RFS0479_036721	Soldier
1387	4.95	35056	RFS0495_035056	Soldier
1388	4.95	36461	RFS0495_036461	Soldier
1389	4.96	38543	RFS0496_038543	Soldier
1390	4.90	41342	RFS0490_041342	Soldier
1391	5.02	45335	RFS0502_045335	Soldier
1392	5.11	46626	RFS0511_046626	Soldier
1393	5.18	45790	RFS0518_045790	Soldier
1394	5.43	39919	RFS0543_039919	Soldier
1395	5.54	40916	RFS0554_040916	Soldier
1396	5.79	41525	RFS0579_041525	Soldier
1397	5.67	42251	RFS0567_042251	Soldier
1398	5.80	43369	RFS0580_043369	Soldier
1399	5.40	43683	RFS0540_043683	Soldier
1400	5.45	46419	RFS0545_046419	Soldier
1401	5.60	46723	RFS0560_046723	Soldier
1402	5.31	49714	RFS0531_049714	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1403	5.38	49594	RFS0538_049594	Soldier
1404	5.40	56970	RFS0540_056970	Soldier
1405	5.40	58822	RFS0540_058822	Soldier
1406	5.26	66379	RFS0526_066379	Soldier
1407	5.39	67774	RFS0539_067774	Soldier
1408	5.48	66217	RFS0548_066217	Soldier
1409	5.62	64033	RFS0562_064033	Soldier
1410	5.59	61604	RFS0559_061604	Soldier
1412	5.80	58679	RFS0580_058679	Soldier
1413	5.79	62346	RFS0579_062346	Soldier
1414	6.06	67400	RFS0606_067400	Soldier
1415	6.11	64556	RFS0611_064556	Soldier
1416	6.12	61628	RFS0612_061628	Soldier
1417	6.12	58799	RFS0612_058799	Soldier
1418	6.15	56058	RFS0615_056058	Soldier
1419	6.21	55913	RFS0621_055913	Soldier
1420	5.89	56902	RFS0589_056902	Soldier
1421	5.96	56187	RFS0596_056187	Soldier
1422	5.94	54690	RFS0594_054690	Soldier
1423	6.00	50420	RFS0600_050420	Soldier
1424	5.93	47572	RFS0593_047572	Soldier
1425	6.00	47735	RFS0600_047735	Soldier
1426	6.22	47863	RFS0622_047863	Soldier
1427	6.11	47259	RFS0611_047259	Soldier
1428	5.99	45119	RFS0599_045119	Soldier
1428	5.99	45119	RFS0599_045119	Soldier
1429	5.98	42662	RFS0598_042662	Soldier
1430	5.98	39790	RFS0598_039790	Soldier
1431	5.97	37922	RFS0597_037922	Soldier
1432	5.94	34886	RFS0594_034886	Soldier
1433	5.81	28352	RFS0581_028352	Soldier
1434	5.82	25487	RFS0582_025487	Soldier
1435	5.59	28103	RFS0559_028103	Soldier
1436	6.37	48237	RFS0637_048237	Soldier
1437	6.40	46576	RFS0640_046576	Soldier
1438	6.59	38735	RFS0659_038735	Soldier
1439	6.89	45005	RFS0689_045005	Soldier
1440	6.76	25516	RFS0620_031750	Soldier
1441	6.94	23676	RFS0694_023676	Soldier
1442	7.03	23283	RFS0703_023283	Soldier
1443	7.44	20558	RFS0744_020558	Soldier
1444	7.43	25961	RFS0743_025961	Soldier
1445	7.73	24206	RFS0773_024206	Soldier
1446	7.84	22156	RFS0784_022156	Soldier
1447	7.86	21033	RFS0786_021033	Soldier
1448	9.43	71407	RFS0943_071407	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1449	9.55	70733	RFS0955_070733	Soldier
1450	9.65	71160	RFS0965_071160	Soldier
1451	7.71	16368	RFS0771_016368	Soldier
1452	6.80	14929	RFS0680_014929	Soldier
1453	6.54	13221	RFS0654_013221	Soldier
1454	6.57	12659	RFS0657_012659	Soldier
1455	6.74	12224	RFS0674_012224	Soldier
1456	6.67	11693	RFS0667_011693	Soldier
1457	6.64	11613	RFS0664_011613	Soldier
1458	6.28	11127	RFS0628_011127	Soldier
1459	6.82	11429	RFS0682_011429	Soldier
1460	7.40	11916	RFS0740_011916	Soldier
1461	7.45	11547	RFS0745_011547	Soldier
1462	7.52	11297	RFS0752_011297	Soldier
1463	7.48	6811	RFS0748_006811	Soldier
1464	7.69	6781	RFS0769_006781	Soldier
1465	6.43	21120	RFS0643_021120	Soldier
1466	6.45	18787	RFS0645_018787	Soldier
1467	6.47	16727	RFS0647_016727	Soldier
1468	6.35	16359	RFS0635_016359	Soldier
1469	6.10	16564	RFS0610_016564	Soldier
1470	5.96	10463	RFS0596_010463	Soldier
1471	5.49	10356	RFS0549_010356	Soldier
1472	5.74	9645	RFS0574_009645	Soldier
1473	5.70	8758	RFS0570_008758	Soldier
1474	5.83	9904	RFS0583_009904	Soldier
1475	5.82	8960	RFS0582_008960	Soldier
1476	5.45	6275	RFS0545_006275	Soldier
1477	5.59	6801	RFS0559_006801	Soldier
1478	5.71	7108	RFS0571_007108	Soldier
1479	6.28	9620	RFS0628_009620	Soldier
1480	6.29	9044	RFS0629_009044	Soldier
1481	6.24	8268	RFS0624_008268	Soldier
1482	6.13	7680	RFS0613_007680	Soldier
1483	6.17	7198	RFS0617_007198	Soldier
1484	6.31	7889	RFS0631_007889	Soldier
1485	6.41	7502	RFS0641_007502	Soldier
1486	6.42	6760	RFS0642_006760	Soldier
1487	6.10	5400	RFS0610_005400	Soldier
1488	5.95	5028	RFS0595_005028	Soldier
1489	4.66	3250	RFS0466_003250	Soldier
1490	4.43	3579	RFS0443_003579	Soldier
1491	4.63	4437	RFS0463_004437	Soldier
1492	4.76	7923	RFS0476_007923	Soldier
1493	4.63	5006	RFS0463_005006	Soldier
1494	4.16	3817	RFS0416_003817	Soldier

Spot Map #	Experimental pI	Experimental MW (Da)	Protein Identifier	Caste
1495	3.86	4202	RFS0386_004202	Soldier
1496	3.72	4543	RFS0372_004543	Soldier
1497	3.72	5878	RFS0372_005878	Soldier
1498	3.81	5621	RFS0381_005621	Soldier
1499	4.05	5612	RFS0405_005612	Soldier
1500	3.99	6494	RFS0399_006494	Soldier
1501	3.46	7710	RFS0346_007710	Soldier
1502	3.56	7702	RFS0356_007702	Soldier
1503	3.70	7512	RFS0370_007512	Soldier
1504	3.71	7981	RFS0371_007981	Soldier
1505	3.63	8884	RFS0363_008884	Soldier
1506	3.33	9309	RFS0333_009309	Soldier
1507	3.40	10006	RFS0340_010006	Soldier
1508	3.57	10077	RFS0357_010077	Soldier
1509	4.10	9786	RFS0410_009786	Soldier
1510	6.24	69052	RFS0624_069052	Soldier
1512	4.67	9582	RFS0467_009582	Soldier
1513	5.46	24546	RFS0546_024546	Soldier
1514	5.64	25192	RFS0564_024192	Soldier
1515	5.73	25531	RFS0573_025531	Soldier
1516	5.55	59715	RFS0555_059715	Soldier
1527	5.21	74284	RFS0521_074284	Soldier
1528	5.19	71905	RFS0519_071905	Soldier
1529	5.26	72499	RFS0526_072499	Soldier
1530	5.25	71260	RFS0525_071260	Soldier
1531	5.41	72873	RFS0541_072873	Soldier
1532	5.19	67341	RFS0519_067341	Soldier
1533	5.28	68359	RFS0528_068359	Soldier
1534	5.38	70198	RFS0538_070198	Soldier
1535	5.48	72423	RFS0548_072423	Soldier
1536	5.48	710710	RFS0548_071710	Soldier
1537	5.59	72327	RFS0559_072327	Soldier
1538	5.65	73641	RFS0565_073641	Soldier
1539	5.66	50794	RFS0566_050794	Soldier
1540	5.66	50794	RFS0566_050794	Soldier

End of Table

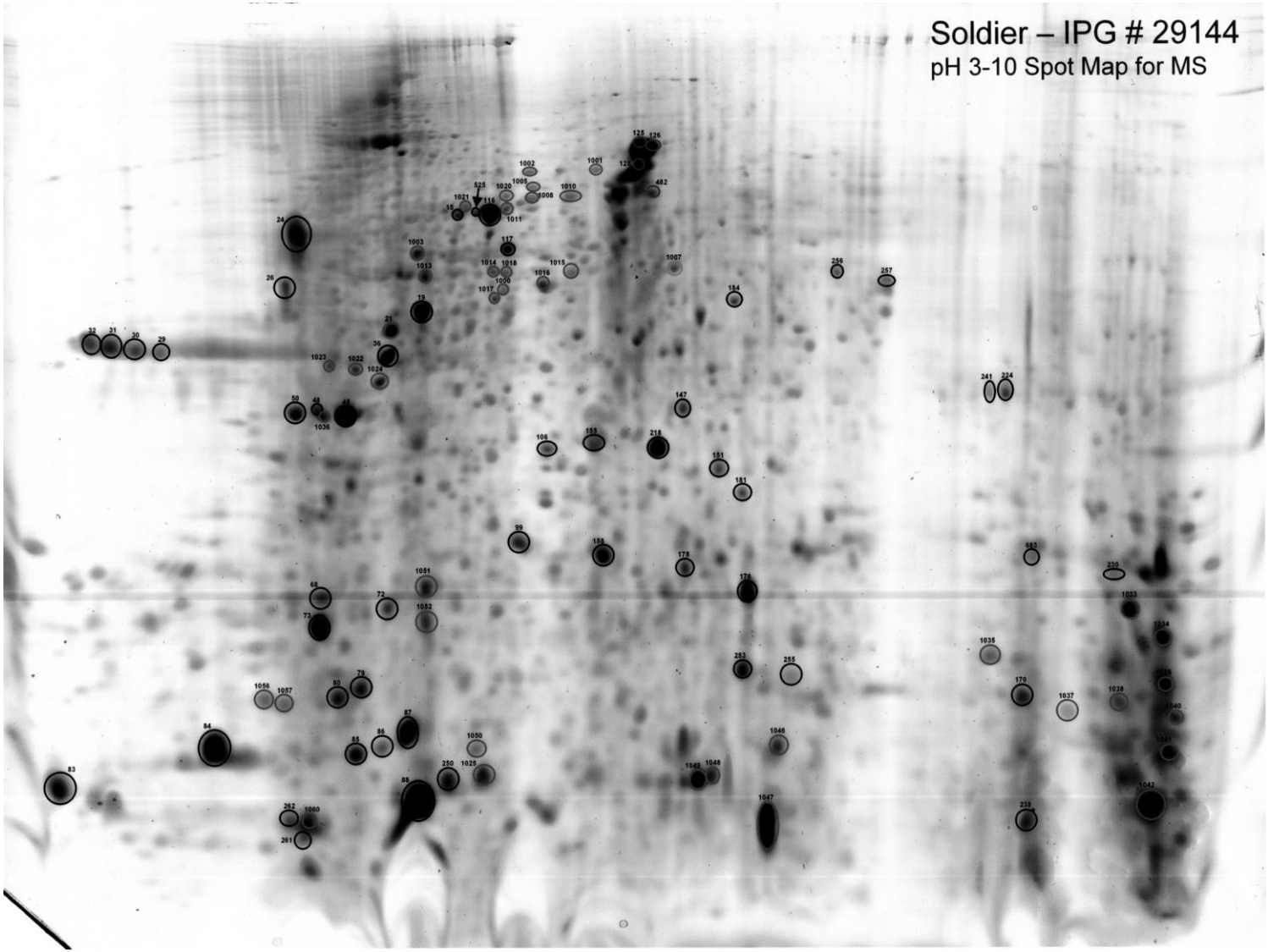


Fig. 45. Soldier caste mass spectrometry reference map for *Reticulitermes flavipes* (3–10 pH range).

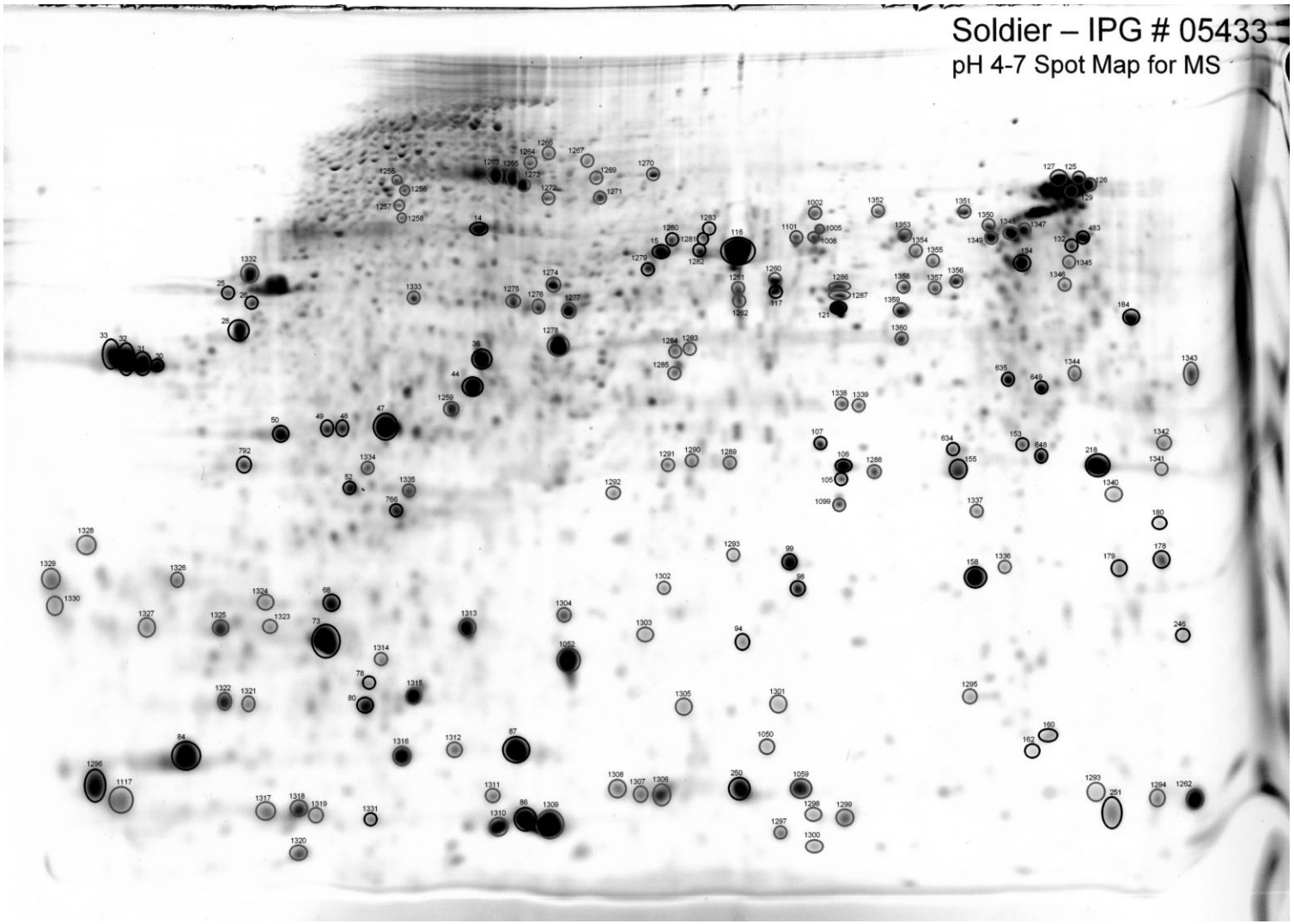


Fig. 46. Soldier caste reference mass spectrometry map for *Reticulitermes flavipes* (4–7 pH range).

Table 7. Putative soldier protein identifications

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	14	PMF	4.81	64668	RFW0481_064668	36	38	CG13861-PA	*Drosophila melanogaster*
good	15	PMF	4.95	56644	RFW0495_056644	58	0.26	exuperantia 1	*Drosophila pseudoobscura*
high	16	PMF	5.35	57285	RFW0535_057285	83	0.00076	Arginine kinase (EC 2.7.3.3)	*Blatella germanica* (German cockroach)
low	19	PMF	5.30	34612	RFW0530_034612	60	0.14	CG11727-PA, isoform A	*Drosophila melanogaster*
moderate	21	PMF	5.11	31950	RFW0511_031950	53	0.78	AY122207 NID	*Drosophila melanogaster*
high	24	PMF	4.49	51451	RFW0449_051451	73	0.0071	Tropomyosin (Fragment)	*Lepisma saccharina* (Silverfish)
low	25	PMF	4.48	39318	RFW0448_039318	51	1.1	GA19596-PA (Fragment)	*Drosophila pseudoobscura*
low	26	PMF	4.56	38734	RFW0456_038734	38	23	GA2122-PA (Fragment)	*Drosophila pseudoobscura*
moderate	28	PMF	4.46	33296	RFW0456_038734	69	0.88	rswcc0_001545.y1 swc Bombyx mori cDNA, mRNA sequence	*Bombyx mori*
low	29	PMF	4.12	34327	RFW0412_034327	43	8.2	Hypothetical protein	*Manduca sexta* (Tobacco hawkmoth, Tobacco hornworm)
low	30	PMF	4.05	29393	RFW0405_029393	43	7	Sp17 protein precursor	*Chironomus tentans* (midge)
low	31	PMF	3.97	29682	RFW0397_039682	51	1.2	Tehao (Fragment)	*Drosophila melanogaster*
low	32	PMF	3.87	29884	RFW0387_029884	51	1.2	Moj29 (Fragment)	*Drosophila mojavensis* (Fruit fly)
low	33	PMF	3.77	29998	RFW0377_029998	45	4.7	ENSANGP00000013890	*Anopheles gambiae* str. PEST
low	36	PMF	5.12	29071	RFW0512_029071	51	1.2	Nucleoporin P54	*Aedes aegypti* (Yellowfever mosquito)
low	44	PMF	5.08	26522	RFW0508_026522	47	3.7	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
high	45	PMF	5.02	24972	RFW0502_024972	66	0.037	GA12823-PA (Fragment)	*Drosophila pseudoobscura*
good	47	PMF	4.94	24245	RFW0494_024245	58	0.21	ENSANGP00000011670 (Fragment)	*Anopheles gambiae* str. PEST
low	48	PMF	4.73	28240	RFW0473_028240	50	1.8	PREDICTED: similar to transcription factor 15	*Apis mellifera*
moderate	49	PMF	4.65	27120	RFW0465_027120	55	0.51	PREDICTED: similar to CG3758-PA	*Tribolium castaneum*
good	50	PMF	4.69	24164	RFW0469_024164	56	0.42	PREDICTED: similar to CG11146-PA	*Tribolium castaneum*

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	52	PMF	4.68	20266	RFW0468_020266	50	1.6	Pheromone binding protein	*Ostrinia nubilalis* (European corn borer)
low	68	PMF	3.41	17625	RFW0467_015450	55	21	EAST77279D BEA Boophila microplus cDNA clone BEABT40, mRNA sequence	*Rhipicephalus microplus*
high	73	PMF	4.61	14460	RFW0461_014460	74	0.0057	DM5CACT2 NID	*Drosophila melanogaster*
low	78	PMF	4.70	13319	RFW0470_013319	39	21	Similar to Drosophila melanogaster CG12338 (Fragment)	*Drosophila yakuba* (Fruit fly)
low	79	PMF	4.99	12602	RFW0575_090695	39	21	DM5CACT2 NID	*Drosophila melanogaster*
low	80	PMF	4.72	12704	RFW0472_012704	43	8.3	putative 60S ribosomal protein L13A	*Diaphorina citri*
low	83	PMF	3.58	10573	RFW0472_012704	47	3.3	CG1965-PA	*Drosophila melanogaster*
moderate	84	PMF	4.15	11457	RFW0415_011457	54	0.67	unnamed protein product	*Drosophila melanogaster*
low	85	PMF	4.79	7216	RFW0479_007216	28	2.1e+002	Peptidyl-prolyl cis-trans isomerase f, ppif.	*Aedes aegypti* (Yellowfever mosquito)
low	86	PMF	4.88	7467	RFW0488_007467	44	7.1	ENSANGP00000014348	*Anopheles gambiae* str. PEST
low	87	PMF	5.00	7326	RFW0500_007326	51	1.5	GA21097-PA	*Drosophila pseudoobscura*
low	88	PMF	5.16	6220	RFW0516_006220	37	32	cupiennin-1d	*Cupiennius salei* (Wandering spider)
low	94	PMF	5.85	12752	RFW0585_012752	50	1.8	PREDICTED: similar to CG9177-PB, isoform B isoform 1	*Tribolium castaneum*
low	98	PMF	6.03	17694	RFW0603_017694	27	3.6e+002	GA1996-PA	*Drosophila pseudoobscura*
moderate	99	PMF	5.99	18483	RFW0599_018483	52	0.94	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
moderate	105	PMF	5.95	23523	RFW0595_023523	55	0.57	hypothetical protein AaeL_AAEL003880	*Aedes aegypti* (Yellowfever mosquito)
moderate	106	PMF	6.07	23895	RFW0607_023895	54	0.66	GH03748p	*Drosophila melanogaster*
good	107	PMF	6.05	25156	RFW0605_025156	56	0.44	GA17449-PA	*Drosophila pseudoobscura*
high	116	PMF	5.17	54053	RFW0517_054053	75	0.0059	arginine kinase	*Periplaneta americana*
high	117	PMF	5.16	45647	RFW0516_045647	66	0.037	AE003652 NID	*Drosophila melanogaster*
low	121	PMF	6.09	50838	RFW0609_050838	47	2.7	CG3085-PA (GH1582p)	*Drosophila melanogaster*
low	125	PMF	6.13	22942	RFW0613_022942	38	24	ENSANGP00000018114 (Fragment)	*Anopheles gambiae* str. PEST

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	126	PMF	5.83	74663	RFW0583_074663	46	4.1	Synaptotagmin, putative	*Aedes aegypti* (Yellowfever mosquito)
low	127	PMF	6.50	76258	RFW0650_076258	48	2.8	Drosophila mauritiana nullo	*Drosophila mauritiana*
low	129	PMF	5.74	71500	RFW0574_071500	47	3	AY084126 NID	*Drosophila melanogaster*
low	132	PMF	6.56	63552	RFW0656_063552	49	1.9	Putative Cyp28c1 protein (Fragment)	*Drosophila yakuba* (Fruit fly)
good	134	PMF	5.41	40788	RFW0541_040788	55	0.48	ENSANGP00000028647 (Fragment)	*Anopheles gambiae* str. PEST
high	151	PMF	6.62	25617	RFW0662_025617	66	1.9	ESG0120a.B21_F06.3prime ESG01 Drosophila melanogaster cDNA 3' similar to CT32334, mRNA sequence	*Drosophila melanogaster*
low	153	PMF	6.18	26145	RFW0618_026145	47	3.2	chiffon CG5813-PA, isoform A	*Drosophila melanogaster*
low	155	PMF	6.13	24721	RFW0613_024721	47	3.4	PREDICTED: similar to Cip4 CG15015-PA	*Apis mellifera*
low	158	PMF	6.30	18089	RFW0630_018089	44	6.3	PREDICTED: similar to Probable cytochrome P450 6g2 (CYPVIG2)	*Apis mellifera*
low	160	PMF	6.44	8270	RFW0644_008270	37	32	PREDICTED: similar to Huntingtin interacting protein K, partial	*Apis mellifera*
low	162	PMF	6.42	7466	RFW0642_007466	47	3.4	PREDICTED: similar to Cip4 CG15015-PA	*Apis mellifera*
low	167	PMF	7.14	9069	RFW0714_009069	51	1.3	Creatine kinase (Arginine or creatine kinase)	*Aedes aegypti* (Yellowfever mosquito)
low	170	PMF	8.84	8488	RFW0884_008488	41	12	HDC10779	*Drosophila melanogaster*
low	176	PMF	7.26	16382	RFW0726_016382	46	3.8	ATP-dependent RNA helicase	*Aedes aegypti* (Yellowfever mosquito)
low	178	PMF	6.84	17635	RFW0684_017635	49	2.4	esc	*Drosophila bipectinata*
low	179	PMF	6.78	16844	RFW0678_016844	50	1.9	olfactory binding protein	*Leucophaea maderae*
low	180	PMF	6.67	20462	RFW0667_020462	42	10	LD10247p	*Drosophila melanogaster*
low	181	PMF	6.85	23023	RFW0685_023023	36	36	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	184	PMF	6.36	33803	RFW0636_033803	44	6.5	CG5964-PA	*Drosophila melanogaster*
low	186	PMF	6.97	43594	RFW0697_043594	43	7	GA13821-PA (Fragment)	*Drosophila melanogaster*
high	187	PMF	7.34	40554	RFW0734_040554	66	0.041	probable transposition protein - fruit fly (Drosophila mauritiana) transposon mariner	*Drosophila mauritiana*

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	218	PMF	6.24	26767	RFW0624_026767	52	1	Muscular protein 20 (Fragment)	*Cicindela theatina*
moderate	231	PMF	9.28	13170	RFW0928_013170	54	0.62	InaD protein	*Calliphora vicina* (Blue blowfly, Calliphora erythrocephala)
low	234	PMF	9.35	26747	RFW0845_029940	47	3	AE003490 NID	*Drosophila melanogaster*
low	241	PMF	8.34	29611	RFW0834_029611	45	4.7	ESANGP00000025736	*Anopheles gambiae* str. PEST
low	246	PMF	7.00	11230	RFW0700_011230	43	9.5	receptor type guanylyl cyclase	*Bombyx mori*
low	250	PMF	5.59	6440	RFW0559_006440	38	27	Tid56 protein	*Drosophila melanogaster*
low	251	PMF	6.65	5927	RFW0665_006927	42	10	ENSANGP00000017318	*Anopheles gambiae* str. PEST
low	261	PMF	4.46	5858	RFW0446_005858	32	94	CG6049-PA, isoform A (Cg6049-pb, isoform b) (LD27763p)	*Drosophila melanogaster*
low	262	PMF	4.40	5958	RFW0440_005958	39	17	HDC07857	*Drosophila melanogaster*
low	483	PMF	5.38	41754	RFW0538_041754	34	58	Myosin heavy chain 2 (Fragment)	*Drosophila melanogaster*
high	498	PMF	5.42	55865	RFW0542_055865	67	0.028	Similar to Drosphila melanogaster Mhc (Fragment)	*Drosophila yakuba* (Fruit fly)
low	500	PMF	5.37	57577	RFW0537_057577	38	23	Ribosomal protein 49 (Fragment)	*Drosophila virilis* (Fruit fly)
low	505	PMF	5.38	53927	RFW0538_053927	36	39	IP08160p (Fragment)	*Drosophila melanogaster*
good	506	PMF	5.34	52029	RFW0534_052029	56	0.41	GA20408-PA (Fragment)	*Drosophila pseudoobscura*
moderate	507	PMF	5.88	61244	RFW0588_061244	53	0.7	D. melanogaster tropomyosin gene 1 constant region, exon 9 (Fragmetn)	*Drosophila melanogaster*
low	525	PMF	5.08	56862	RFW0508_056862	50	1.3	AF233355NID	*Callinectes sapidus*
low	592	PMF	5.71	47700	RFW0571_047700	42	9.4	sex-specific storage protein 1 precursor - silkworm	*Bombyx mori*
good	593	PMF	5.28	42471	RFW0528_042471	60	0.15	IP06843p (Fragment)	*Drosophila melanogaster*
low	634	PMF	6.15	26122	RFW0615_026122	43	9.7	PREDICTED: similar to CG7175-PA	*Tribolium castaneum*
low	635	PMF	6.03	29396	RFW0603_029396	51	1.3	GA10231-PA	*Drosophila pseudoobscura*
low	648	PMF	6.18	27299	RFW0618_027299	48	3	famiily 4 cytochrome P450	*Coptotermes acinaciformis*
good	649	PMF	6.51	20946	RFW0651_020946	59	0.21	DNA-binding protein smubp-2	*Aedes aegypti* (Yellowfever mosquito)

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	684	PMF	8.77	14631	RFW0877_014631	41	11	GA16567-PA (Fragment)	*Drosophila pseudoobscura*
low	766	PMF	4.95	19057	RFW0495_019057	41	11	Est (Fragment)	*Culex pipiens* (House mosquito)
low	792	PMF	4.26	21523	RFW0426_021523	45	4.7	ENSANGP00000003954	*Anopheles gambiae* str. PEST
high	1033	PMF	9.41	16088	RFS0941_016088	67	0.027	Mircrotubule-associated protein	*Aedes aegypti* (Yellowfever mosquito)
low	1034	PMF	9.63	14651	RFS0963_014651	49	1.8	AF163664 NID	*Drosophila melanogaster*
low	1037	PMF	9.02	10741	RFS0902_010741	47	3	Wingless (Fragment)	*Pleistodontes froggatti*
low	1039	PMF	9.59	12237	RFS0959_012237	46	3.4	ENSANGP00000006025 (Fragment)	*Anopheles gambiae* str. PEST
low	1040	PMF	6.94	10567	RFS0694_010567	23	7.8e+002	ENSANGP00000024489	*Anopheles gambiae* str. PEST
low	1041	PMF	9.62	9970	RFS0962_009970	49	1.9	RNA m5u methyltransferase	*Aedes aegypti* (Yellowfever mosquito)
low	1045	PMF	5.20	12379	RFS0520_012379	45	4.5	Amylase (Fragment)	*Drosophila rufa*
low	1050	PMF	5.54	6007	RFS0554_006007	43	8.1	ENSANGP00000020737	*Anopheles gambiae* str. PEST
low	1052	PMF	5.49	10356	RFS0549_010356	5.9	5.9	PREDICTED: similar to CG30010-PA	*Tribolium castaneum*
low	1056	PMF	4.50	6075	RFS0450_006075	21	1.2e+003	GA10255-PA (Fragment)	*Drosophila pseudoobscura*
low	1059	PMF	5.92	5631	RFS0592_005631	32	1e+002	elongation factor-1 alpha	*Cymbomorpha sp.*
low	1060	PMF	5.53	12716	RFS0553_012716	27	301e+002	ENSANGP00000026057 (Fragment)	*Anopheles gambiae* str. PEST
moderate	1061	PMF	5.24	49485	RFS0524_049485	52	0.92	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1062	PMF	5.28	45227	RFS0528_045227	45	4.6	ENSANGP00000024695	*Anopheles gambiae* str. PEST
low	1077	PMF	4.74	34154	RFS0474_034154	22	1e+003	RE03722p	*Drosophila melanogaster*
good	1078	PMF	4.83	32546	RFS0483_032546	55	0.46	CG4937-PA	*Drosophila melanogaster*
low	1099	PMF	5.99	18445	RFS0599_018445	40	17	centrin	*Bombyx mori*
low	1117	PMF	3.75	5631	RFS0375_005631	44	7.4	CG10026-PA, isoform A	*Drosophila melanogaster*
low	1124	PMF	7.30	7325	RFS0730_007325	35	52	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1133	PMF	9.30	11080	RFS0930_011080	47	2.9	IP12707p	*Drosophila melanogaster*
high	1180	PMF	6.70	47754	RFS0575_036535	64	0.062	Pheromone binding protein (Fragment)	*Ostrinia nubilalis* (European corn borer)

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	1215	PMF	5.37	14014	RFS0537_014014	44	5.4	Ubiquitin ligase E3 (Fragment)	*Aedes aegypti* (Yellowfever mosquito)
low	1216	PMF	5.38	11466	RFS0538_011466	46	3.9	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1221	PMF	7.41	10268	RFS0741_010268	41	12	Myosin heavy chain 2 (Fragment)	*Drosophila melanogaster*
low	1223	PMF	7.27	10173	RFS0727_010173	42	10	HDC19173	*Drosophila melanogaster*
low	1227	PMF	7.31	8888	RFS0731_008888	35	51	Triose phosphate isomerase (Fragment)	*Drosophila kikkawai*
low	1234	PMF	6.97	8391	RFS0697_008391	31	1.1e+002	ML domain-containing protein (Fragment)	*Ixodes ricinus* (Sheep tick)
low	1247	PMF	8.53	11966	RFS0853_011966	46	3.7	ENSANGP00000014611 (Fragment)	*Anopheles gambiae* str. PEST
low	1253	PMF	9.52	7823	RFS0952_007823	51	1.2	ADP/ATP translocase	*Lucilia cuprina* (Green bottlefly, Australian sheep blowfly)
low	1255	PMF	4.73	62749	RFS0473_062749	32	86	CG10622-PB, isoform B	*Drosophila melanogaster*
high	1256	PMF	4.68	54868	RFS0468_054868	73	0.008	Myosin heavy chain 4, muscle (Fragment)	*Drosophila melanogaster*
low	1257	PMF	4.69	57069	RFS0469_057069	40	16	triosephosphate isomerase (EC 5.3.1.1) (Fragment)	*Calliphora vicina* (Blue blowfly, Calliphora erythrocephala)
low	1258	PMF	4.68	54868	RFS0468_054868	130	1.5e-008	Myosin heavy chain, nonmuscle or smooth muscle	*Aedes aegypti* (Yellowfever mosquito)
low	1259	PMF	4.70	29745	RFS0470_029745	48	2.9	Nipped-A	*Drosophila melanogaster*
low	1260	PMF	5.31	49981	RFS0531_049981	43	8	CG33060-PA	*Drosophila melanogaster*
low	1261	PMF	7.28	45603	RFS0728_045603	48	2.4	Arginase	*Aedes aegypti* (Yellowfever mosquito)
low	1263	PMF	4.90	78351	RFS0490_078351	42	8.6	CG11990-PA (Hyrax) (LD47989p)	*Drosophila melanogaster*
low	1264	PMF	4.97	94994	RFS0497_094994	47	2.7	Myosin heavy chain (Fragment)	*Drosophila melanogaster*
low	1265	PMF	4.92	79680	RFS0492_079680	50	1.3	CG11156 (Fragment)	*Drosophila simulans*
low	1266	PMF	5.02	81739	RFS0502_081739	33	78	Putative 22.5kDa secreted protein	*Ixodes scapularis* (Black-legged tick) (Deer tick)
moderate	1267	PMF	5.07	79888	RFS0507_079888	54	0.64	Myosin heavy chain 3, muscle (Fragment)	*Drosophila melanogaster*
high	1269	PMF	5.06	85569	RFS0506_085569	65	0.043	AE003652 NID	*Drosophila melanogaster*
low	1270	PMF	5.19	75459	RFS0519_075459	51	1.1	AE003652 NID	*Drosophila melanogaster*

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
high	1271	PMF	5.02	75350	RFS0502_075350	69	0.017	Myosin heavy chain 4, muscle (Fragment)	*Drosophila melanogaster*
low	1272	PMF	4.94	65304	RFS0494_065304	37	28	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
good	1273	PMF	4.93	77136	RFS0493_077136	60	0.15	Cytochrome P450	*Aedes aegypti* (Yellowfever mosquito)
high	1274	PMF	4.82	45011	RFS0482_045011	82	0.00092	beta-tubulin	*Bombyx mori*
moderate	1275	PMF	4.74	41957	RFS0474_041957	54	0.74	PREDICTED: similar to outspread CG3479-PA, isoform A	*Apis mellifera*
good	1276	PMF	4.80	41907	RFS0480_041907	58	0.26	Chain, Heat Shock Transcription Factor (Nmr, Restrained Minimized Average Structure)	*Drosophila melanogaster*
moderate	1277	PMF	4.87	41942	RFS0487_041942	54	0.71	IP13307p	*Drosophila melanogaster*
moderate	1278	PMF	4.87	37818	RFS0487_037818	52	0.96	GA17852-PA (Fragment)	*Drosophila pseudoobscura*
low	1279	PMF	5.00	49040	RFS0500_049040	33	75	Triose phosphate isomerase (EC 5.3.1.1) (Fragment)	*Drosophila heteroneura* (Fruit fly)
low	1281	PMF	5.14	59222	RFS0514_059222	49	1.8	Tropomyosin - migatory locust	*Locusta migratoria*
low	1282	PMF	5.11	52891	RFS0511_052891	44	6.5	CG30022-PA	*Drosophila melanogaster*
moderate	1283	PMF	5.18	41114	RFS0518_041114	52	0.92	Protease m1 zinc metalloprotease	*Aedes aegypti* (Yellowfever mosquito)
low	1284	PMF	5.14	39563	RFS0514_039563	45	5.2	Peptidyl-prolyl cis-trans isomerase f, ppif.	*Aedes aegypti* (Yellowfever mosquito)
low	1285	PMF	5.16	37388	RFS0516_037388	46	3.6	Similar to Drosophila melanogaster Mhc (Fragment).	*Drosophila yakuba* (Fruit fly)
high	1287	PMF	5.51	45963	RFS0551_045963	70	0.015	CG11508-PA, isoform A (Cg11508-pb, isoform B) (LD18062p)	*Drosophila melanogaster*
good	1288	PMF	5.50	46713	RFS0550_046713	57	0.28	Ras-related protein, putative	*Aedes aegypti* (Yellowfever mosquito)
low	1289	PMF	5.44	30489	RFS0544_030489	49	2	inaD	*Calliphora vicina* (Blue blowfly, Calliphora erythrocephala)
low	1291	PMF	5.28	29129	RFS0528_029129	49	2.3	TPA: TPA_inf: HDC19521	*Drosophila melanogaster*
low	1292	PMF	5.17	26047	RFS0517_026047	47	3.5	putative transposase yabusame-W	*Bombyx mori*
low	1293	PMF	5.58	23913	RFS0558_023913	38	29	circadian clock protein period	*Sesamia nonagriodes*

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
moderate	1294	PMF	6.76	6093	RFS0676_006093	53	0.93	ENSANGP00000020478	*Anopheles gambiae* str. PEST
low	1295	PMF	6.19	9534	RFS0619_009534	50	1.9	Myosin heavy chain 2 (Fragment)	*Drosophila melanogaster*
low	1297	PMF	5.85	5413	RFS0585_005413	42	12	PREDICTED: similar to Isocitrate dehydrogenase CG7176-PC, isoform C isoform 2	*Apis mellifera*
low	1298	PMF	5.93	6975	RFS0593_006975	35	61	PIWI	*Aedes aegypti* (Yellowfever mosquito)
low	1299	PMF	6.01	6851	RFS0601_006851	42	11	SD14927p	*Drosophila melanogaster*
low	1300	PMF	5.93	5121	RFS0593_005121	36	41	ENSANG00000029123	*Anopheles gambiae* str. PEST
low	1301	PMF	5.75	9762	RFS0575_009762	38	26	malate synthase	*Aedes aegypti* (Yellowfever mosquito)
good	1302	PMF	5.35	15710	RFS0535_015710	57	0.33	vitellogenin	*Pandalus hypsinotus*
low	1303	PMF	5.36	13312	RFS0536_013312	46	4.6	ENSANGP00000029738	*Anopheles gambiae* str. PEST
low	1304	PMF	5.14	14793	RFS0514_014793	47	3.6	CG6744-PA	*Drosophila melanogaster*
low	1305	PMF	5.53	9982	RFS0553_009982	47	3.2	inaD	*Calliphora vicina* (Blue blowfly, Calliphora erythrocephala)
low	1306	PMF	5.31	8070	RFS0531_008070	31	1.4e+002	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1307	PMF	5.24	8222	RFS0524_008222	36	42	lectin-related protein	*Glyptapanteles indiensis*
low	1308	PMF	5.17	8364	RFS0517_008364	49	2.2	GA10772-PA	*Drosophila pseudoobscura*
low	1309	PMF	5.30	5592	RFS0530_005592	36	43	Amyrel	*Drosophila americana texana*
low	1310	PMF	5.17	5578	RFS0517_005578	37	36	ENSANGP00000011961	*Anopheles gambiae* str. PEST
low	1311	PMF	5.16	6301	RFS0516_006301	42	10	conserved hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1312	PMF	4.77	9769	RFS0477_009769	49	2.2	Chain A, 1.0 A Crystal Structure of D129aL130A MUTANT OF Nitrophorin 4 Complexed With Nitric Oxide	*Rhodnius prolixus*
moderate	1313	PMF	4.94	14333	RFS0494_014333	53	0.85	actin E2	*Drosophila americana*
low	1314	PMF	4.74	13222	RFS0474_013222	45	5.1	Low molecular weight protein-tyrosine-phosphatase	*Aedes aegypti* (Yellowfever mosquito)

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
moderate	1315	PMF	4.76	11756	RFS0476_011756	54	0.63	actin E2	*Drosophila americana*
low	1316	PMF	4.59	9854	RFS0459_009854	42	12	ENSANGP00000025780	*Anopheles gambiae* str. PEST
low	1317	PMF	5.24	5874	RFS0524_005874	43	9.3	prefoldin, subunit, putative	*Aedes aegypti* (Yellowfever mosquito)
low	1318	PMF	4.68	5843	RFS0468_005843	31	1.5e+002	GA14787-PA	*Drosophila pseudoobscura*
low	1319	PMF	4.74	5805	RFS0474_005805	34	76	PREDICTED: similar to mitogen-activated protein kinase kinase 7	*Apis mellifera*
low	1320	PMF	4.69	4986	RFS0469_004986	33	97	PREDICTED: similar to Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial precursor	*Apis mellifera*
low	1321	PMF	4.23	12311	RFS0423_012311	44	6.6	PREDICTED: similar to CG11146-PA	*Tribolium castaneum*
low	1322	PMF	4.13	12564	RFS0413_012564	39	24	mRNA transport regulator 3	*Bombyx mori*
low	1323	PMF	4.39	14907	RFS0439_014907	34	62	CG8478-PA, isoform A	*Drosophila melanogaster*
low	1324	PMF	4.36	16062	RFS0436_016062	15	4.6e+003	CG31415-PA (IP07196p)	*Drosophila melanogaster*
low	1325	PMF	4.16	15149	RFS0416_015149	63	3.3	CS_gil_24F01_M13Reverse Blue crab gill, normalized Callinectes sapidus cDNA clone CS_gil_24F01 5'	*Callinectes sapidus*
good	1326	PMF	4.00	17141	RFS0400_017141	62	0.11	vacuolar ATP synthase subunit H	*Bombyx mori*
low	1327	PMF	3.87	15354	RFS0387_015354	58	9.9	EST774852 BEA Boophilus microplus cDNA clone BEAC530, mRNA sequence	*Rhipicephalus microplus*
high	1328	PMF	4.18	17213	RFS0418_017213	66	0.048	RNA polymerase II largest aubunit	*Ctenolepisma lineata*
high	1329	PMF	4.10	14939	RFS0410_014939	80	0.0017	Chain, Calmodulin	*Drosophila melanogaster*
moderate	1330	PMF	4.11	13436	RFS0411_013436	55	0.51	ENSANGP00000002593	*Anopheles gambiae* str. PEST
moderate	1331	PMF	4.87	5724	RFS0487_005724	53	0.97	Peroxiredoxin 6005 CG3083-PA	*Drosophila melanogaster*
low	1332	PMF	4.57	50612	RFS0457_050612	50	1.5	Condensin	*Aedes aegypti* (Yellowfever mosquito)
low	1333	PMF	4.97	46237	RFS0497_046237	36	37	ENSANGP00000006196 (Fragment)	*Anopheles gambiae* str. PEST

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
low	1334	PMF	4.56	24137	RFS0456_024137	32	92	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1335	PMF	4.68	23173	RFS0468_023173	39	19	Cytochrome P450	*Aedes aegypti* (Yellowfever mosquito)
low	1336	PMF	6.16	15009	RFS0616_015009	29	2e+002	CG14612-PA	*Drosophila melanogaster*
low	1337	PMF	6.05	18579	RFS0605_018579	39	19	ENSANGP00000020301	*Anopheles gambiae* str. PEST
low	1338	PMF	5.28	29506	RFS0528_029506	25	5.2e+002	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1339	PMF	5.63	29160	RFS0563_029160	36	40	ENSANGP00000015927 (Fragment)	*Anopheles gambiae* str. PEST
low	1340	PMF	6.44	18343	RFS0644_018343	38	25	CG5310-PA (Type 6 nucleoside diphosphate kinase)	*Drosophila melanogaster*
low	1341	PMF	6.77	22953	RFS0677_022953	36	34	AE003828 NID	*Drosophila melanogaster*
low	1342	PMF	6.77	25638	RFS0677_025638	38	22	ENSANGP00000025572	*Anopheles gambiae* str. PEST
low	1343	PMF	6.84	33522	RFS0684_033522	42	10	CM000070 NID	*Drosophila pseudoobscura*
high	1344	PMF	6.25	28321	RFS0625_028321	67	0.031	GA18828-PA	*Drosophila pseudoobscura*
low	1345	PMF	6.54	52642	RFS0654_052642	99	99	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1346	PMF	6.54	47834	RFS0654_047834	37	30	ENSANGP00000020952 (Fragment)	*Anopheles gambiae* str. PEST
low	1347	PMF	6.45	59428	RFS0645_059428	37	31	Ribosomal protein L32 (Fragment)	*Drosophila willistoni*
low	1348	PMF	6.41	58587	RFS0641_058587	33	80	CG11928-PA	*Drosophila melanogaster*
good	1349	PMF	6.36	58473	RFS0636_058473	63	0.076	Heat shock transcription factor (nmr, restrained minimized average structure)	*Drosophila melanogaster*
moderate	1350	PMF	6.35	61006	RFS0635_061006	55	0.52	ENSANGP00000029378 (Fragment)	*Anopheles gambiae str. PEST*
low	1351	PMF	5.16	61234	RFS0516_061234	36	39	IKKgamma (Fragment)	*Drosophila simulans*
high	1352	PMF	6.09	64784	RFS0609_064784	85	0.00044	Similar to Drosophila melanogaster Mhc (Fragment).	*Drosophila yakuba* (Fruit fly)
low	1353	PMF	6.60	6636	RFS0660_006636	39	24	stathmin CG31641-PA, isoform A	*Drosophila melanogaster*
moderate	1354	PMF	6.17	54588	RFS0617_054588	53	0.68	AF233355 NID	*Callinectes sapidus*
low	1355	PMF	6.22	52820	RFS0622_052820	46	3.6	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)

Confidence	Map Spot #	MS Type	Experimental pI	Experimental Mass	Protein Identifier	Score	Expect	Description	Source
good	1356	PMF	6.28	49197	RFS0628_049197	58	0.25	Hypothetical protein (Fragment)	*Aedes aegypti* (Yellowfever mosquito)
low	1357	PMF	5.76	42231	RFS0576_042231	39	20	Synaptose-associated protein of 25kDa	*Procambarus clarkii* (Red swamp crayfish)
low	1358	PMF	5.67	43247	RFS0567_043247	44	6.5	PLU	*Drosophila yakuba* (Fruit fly)
moderate	1359	PMF	5.69	39750	RFS0569_039750	52	0.96	Putative arginine kinase	*Homalodisca coagulata* (Glassy-winged sharpshooter)
high	1360	PMF	5.70	36013	RFS0570_036013	76	0.0041	AE003746 NID	*Drosophila melanogaster*
high	1397	PMF	5.67	42251	RFS0567_042251	66	0.034	Arginine kinase (EC 2.7.3.3)	*Blatella germanica* (German cockroach)
low	1398	PMF	5.80	43369	RFS0580_043369	35	49	aptotoxin VI	*trap-door spider* (Aptostichus schlinger)
low	1412	PMF	5.80	58679	RFS0580_058679	49	2	IP12707p	*Drosophila melanogaster*
low	1413	PMF	5.79	62346	RFS0579_062346	45	4.7	GA20161-PA (Fragment)	*Drosophila melanogaster*
moderate	1416	PMF	6.12	61628	RFS0612_061628	52	1	myosin heavy chain, thorax-specific fruit fly (Fragment)	*Drosophila melanogaster*
high	1426	PMF	6.22	47863	RFS0622_047863	73	0.0075	Putative arginine kinase	*Homalodisca coagulata* (Glassy-winged sharpshooter)
low	1427	PMF	6.11	47259	RFS0611_047259	31	1.1e+002	AT31792p	*Drosophila melanogaster*
low	1440	PMF	6.76	25516	RFS0620_031750	51	1.2	ENSANGP00000020733 (Fragment)	*Anopheles gambiae* str. PEST
low	1510	PMF	6.24	69052	RFS0624_069052	45	4.2	Hypothetical protein	*Aedes aegypti* (Yellowfever mosquito)
low	1515	PMF	5.73	25531	RFS0573_025531	42	10	Molybdenum cofactor synthesis-step 1 protein A splie type I	*Drosophila melanogaster*
low	1516	PMF	5.55	59715	RFS0555_059715	42	9.9	AE003814 NID	*Drosophila melanogaster*
low	1539	PMF	5.66	50794	RFS0566_050794	27	3.7e+002	TPA: TPA_inf: HDC14429	*Drosophila melanogaster*
low	1540	PMF	5.66	50794	RFS0566_050794	48	2.2	Cytochrome P450 (Fragment)	*Blatella germanica* (German cockroach)

End of Table

SUMMARY AND CONCLUSIONS

Termites are a social insect with an important economic impact in the United States of America and the world. Much of termite research is geared toward development of potential control methods. Yet termites remain open to discovery from the aspect of fundamental science.

Termites from three separate colonies were studied, one from Stillwater, OK, and two from the Nature Conservancy's Tallgrass Prairie Preserve near Pawhuska, in north-east OK. Gel replicates of proteins were produced from whole-body extracts generated from workers of each colony and were differentially compared using Dymension gel-analysis software. Additionally, one gel from each colony was manually compared with a gel from each other colony. These gel comparisons demonstrated less than 10% of differentially expressed proteins among the colonies. Overall, gels had striking similarities that could prove useful as taxonomic indicators. As only a single species was observed, further protein comparisons with other species will be required prior to any conclusion of the taxonomic value.

Of the protein differences among the colony comparison gels, the tallgrass prairie gels were the most similar to each other and the Stillwater gel was the least similar compared with the tallgrass prairie gels. This indicates geographical factors may influence protein profiles. Geographical differences could include factors such as soil composition, available food resources, seasonal temperature range, or rainfall to name a

few. Additional study will be required to identify the influence of varied geographical factors on differentially expressed proteins.

Gel replicates from soldiers from the Stillwater colony were compared with the Colony 1 worker gels. These comparisons indicated substantial differences between the worker and soldier protein profiles. There are several possible explanations for variation between the different castes' protein profiles. Workers and soldiers exhibit visible morphological differences. Thus, different structural proteins may be present. Another possibility may be non-discovered differences in physiology. For example, soldier termites do not feed directly on wood, but are fed by workers through tropholaxis. Physiological differences may have developed to compensate for differences in feeding behavior between castes. Mass spectrometry performed on soldier proteins failed to yield high confidence putative identifications of differentially expressed structural proteins. Due to the large quantity of differentially expressed proteins, many proteins remain non-identified within the scope of this study.

Mass spectrometry was undertaken on 310 worker proteins and an additional 142 differentially expressed soldier proteins. Protein samples were processed primarily using MALDI-TOF mass spectrometry. Additionally, mass spectra were separately generated using tandem mass spectrometry (MS/MS). The hexamerin protein group was most commonly identified from a termite protein source. Overall, database searches yielded high-confidence putative identifications for 18.7% of the proteins processed. Comparison between MALDI-TOF yields and MS/MS yields revealed a five-fold increase in putative identifications from 7% to 34%, respectively. This indicates tandem mass spectromtry should be the preferred method for generating putative identifications.

However, equipment to carry out this type of analysis is not available at Oklahoma State University. Another limitation in characterizing the termite proteome is that genome data are lacking for *Reticulitermes flavipes* and as well as other termite genera. Thus, database searches were primarily limited to cross-species identifications based on conserved genes. Correlation of putatively identified proteins with comparable functions in known genes was relatively successful. Most of the proteins would be matched with known related function(s). Examples of the most common 'first-level' functions were 'binding', 'catalytic activity', and 'structural molecule activity'.

Termites are considered primitive insects, but most available insect protein information is based on more advanced insects such as *Drosophila* sp., *Anopheles* sp., and *Culex* sp. It is believed there are substantial genetic differences among insect orders, comparable to differences among plant classes. This possible lack of conserved genes among insect orders may have contributed to the low percentage of significant cross-species protein identifications in this research.

This research establishes a system for analyzing the termite proteome as well as baseline putative protein profiles to facilitate further exploration of termite proteomics. The reference maps generated by this research could be used to explore how various termiticides affect protein expression, or proteins affected by insect growth regulators. Differentially expressed proteins among species or genera could also be identified using the reference maps. Such maps could potentially develop into taxonomic use based on the termite proteome. However, other studies could also follow. For example, the procedures developed by this research could facilitate elucidating differentially expressed proteins between instars or observing changes in protein expression as a termite develops

into its terminal form. Another possibility could be to observe differentially expressed proteins in response to climatic extremes, environmental changes, stress, or differences in food resources.

Establishment of an *R. flavipes* protein database will facilitate comparison among termite species, and among different colonies of the same species, as well as during 'time-point' experiments within the same colony. Differential analysis of the protein maps will continue beyond the current research as will the identification of amino acid sequence tags. Data collected from this study will be archived and periodically compared against protein databases to allow continuing elucidation of putative protein identifications. Additionally, availability of a termite genome will increase the number and confidence of existing and future protein identifications.

LITERATURE CITED

Adam, P. J., R. Boyd, K. L. Tyson, G. C. Fletcher, A. Stamps, L. Hudson, H. R. Poyser, N. Redpath, M. Griffiths, G. Steers, A. L. Harris, S. Patel, J. Berry, J. A. Loader, R. R. Townsend, L. Daviet, P. Legrain, R. Parekh, and J. Terrett. 2003. Comprehensive proteomic analysis of breast cancer cell membranes reveals unique proteins with potential roles in clinical cancer. Journal of Biological Chemistry 278: 6482-6489.

Alban, A., S. O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, and I. Currie. 2003. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 3: 36-44.

Alonso, J., and J. F. Santaren. 2006. Characterization of the *Drosophila melanogaster* ribosomal proteome. Journal of Proteome Research 5: 2025-2032.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Amersham Biosciences. 1998. 2-D Electrophoresis: A Comparison of Carrier Ampholyte and Immobilized pH Gradients.

Anonymous. 1998. Guide to Isoelectric Focusing. Amersham Biosciences. 29 Pp.

Appel, R. D., and D. F. Hochstrasser. 1999. Computer analysis of 2-D images, pp. 363-381. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, N. J.

Archakov, A. I., V. M. Govorun, A. V. Dubanov, Y. D. Ivanov, A. V. Veselovsky, P. Lewi, and P. Janssen. 2003. Protein-protein interactions as a target for drugs in proteomics. Proteomics 3: 380-391.

Austin, J. W., A. L. Szalanski, and B. M. Kard. 2004. Distribution and genetic variation of *Reticulitermes* (Isoptera: Rhinotermitidae) in Oklahoma. Florida Entomologist 87: 152-158.

Back, J. W., V. Notenboom, L. J. de Koning, A. O. Muijsers, T. K. Sixma, C. G. de Koster, and L. de Jong. 2002. Identification of cross-linked peptides for protein interaction studies using mass spectrometry and ^{18}O labeling. Analytical Chemistry 2002: 4417-4422.

Bennett, G. W., J. M. Owens, and R. M. Corrigan [eds.]. 1988. Truman's Scientific Guide To Pest Control Operations. Harcourt Brace Jovanovich, Inc., Duluth, MN. 495 Pp.

Berkelman, T., and T. Stenstedt. 1998. 2-D Electrophoresis. Using Immobilized pH Gradients: Principles and Methods. Amersham Pharmacia manual. 53 Pp.

Bertucci, F., D. Birnbaum, and A. Goncalves. 2006. Proteomics of breast cancer: principles and potential clinical applications. Molecular and Cellular Proteomics 5: 1772-1786.

Bevan, M., I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, R. Bergkamp, W. Dirkse, M. Van Staveren, W. Stiekema, L. Drost, P. Ridley, S. A. Hudson, K. Patel, G. Murphy, P. Piffanelli, H. Wedler, E. Wedler, R. Wambutt, T. Weitzenegger, T. M. Pohl, N. Terryn, J. Gielen, R. Villarroel, R. De Clerck, M. Van Montagu, A. Lecharny, S. Auborg, I. Gy, M. Kreis, N. Lao, T. Kavanagh, S. Hempel, P. Kotter, K. D. Entian, M. Rieger, M. Schaeffer, B. Funk, S. Mueller-Auer, M. Silvey, R. James, A. Montfort, A. Pons, P. Puigdomenech, A. Douka, E. Voukelatou, D. Milioni, P. Hatzopoulos, E. Piravandi, B. Obermaier, H. Hilbert, A. Dusterhoft, T. Moores, J. D. G. Jones, T. Eneva, K. Palme, V. Benes, S. Rechman, W. Ansorge, R. Cooke, C. Berger, M. Delseny, M. Voet, G. Volckaert, H. W. Mewes, S. Klosterman, C. Schueller, and N. Chalwatzis. 1998. Analysis of 1.9Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 391: 485-488.

Bienvenut, W. V., C. Deon, C. Pasquarello, J. M. Campbell, J.-C. Sanchez, M. L. Vestal, and D. F. Hochstrasser. 2002. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. Proteomics 2: 868-876.

Bignell, D. E. 2000. Introduction to symbiosis, pp. 189-208. *In*: T. Abe, D. E. Bignell and M. Higashi [eds.], Termites: Evolution, Sociality, Symbioses, Ecology. Kluwer Academic Publishers, Boston, MA.

Bio-Rad. 1990. New crosslinker beats bis-acrylamide. European Journal of Biochemistry 189: 461.

Bio-Rad. 2002. 2-D Electrophoresis for Proteomics: A Methods and Product Manual. 53 Pp.

Blum, H., H. Beier, and H. J. Gross. 1987. Modified silver stain procedure. Electrophoresis 8: 93-99.

Bodovitz, S., and T. Joos. 2004. The Proteomics Bottleneck: Strategies for Preliminary Validation of Potential Biomarkers and Drug Targets. Trends in Biotechnology 22: 4-7.

Boonmee, S., K. Imtawil, C. Wongkham, and S. Wongkham. 2003. Comparative proteomic analysis of juvenile and adult liver fluke, *Opisthorchis viverrini*. Acta Tropica 88: 233-238.

Boyer, R. F. 2002. Concepts in Biochemistry. John Wiley & Sons, Inc., New York. 626 Pp.

Braun, R. P., and G. R. Wyatt. 1996. Sequence of the hexameric juvenile hormone-binding protein from the hemolymph of *Locusta migratoria*. The Journal of Biological Chemistry 271: 31756-31762.

Brooksbank, C. 2000. Proteomics: this way up and handle with care. Nature Reviews: Molecular Cell Biology 1: 4.

Brown, K. S., B. M. Kard, and M. P. Doss. 2004. 2002 Oklahoma termite survey (Isoptera). Journal of the Kansas Entomological Society 77: 1-9.

Bruening, G., R. Criddle, J. Preiss, and F. Rudert. 1970. Biochemical Experiments. Wiley-Interscience, New York. 319 Pp.

Bukowska, A., U. Lendeckel, T. Kahne, and A. Goette. 2004. Proteomics in myocardial diseases. Pathology - Research and Practice 200: 135-145.

Burmester, T., and K. Scheller. 1999. Ligands and receptors: common theme in insect storage protein transport. Naturwissenschaften 86: 468-474.

Camaschella, C., A. Roetto, A. Cali, M. De Gobbi, G. Garozzo, M. Carella, N. Majorano, A. Totaro, and P. Gasparini. 2000. The gene TFR2 is mutated in a new type of haemochromatosis mapping to 7q22. Nature Genetics 25: 14-15.

Cash, P. 1998. Characterisation of bacterial proteomes by two-dimensional electrophoresis. Analytica Chimica Acta 372: 121-145.

Celis, J. E., P. Celis, M. Ostergaard, B. Basse, J. B. Lauridsen, G. Ratz, H. H. Rasmussen, T. F. Orntoft, B. Hein, H. Wolf, and A. Celis. 1999. Proteomics and immunohistochemistry define some of the steps involved in the squamous differentiation of the bladder transitional epithelium: A novel strategy for identifying metaplastic lesions. Cancer Research 59: 3003-3009.

Chen, P., S. Nie, W. Mi, X.-C. Wang, and S.-P. Liang. 2004. *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. Rapid Communications in Mass Spectrometry 18: 191-198.

Churchill, G. A. 2002. Fundamentals of experimental design for cDNA microarrays. Nature Genetics 32: 490-495.

Corthals, G. L., V. C. Wasinger, D. F. Hochstrasser, and J.-C. Sanchez. 2000. The dynamic range of protein expression: a challange for proteomic research. Electrophoresis 21: 1104-1115.

Courchesne, P. L., and S. D. Patterson. 1999. Identification of proteins by matrix-assisted laser desorption/ionization mass spectrometry using peptide and fragment ion masses, pp. 487-511. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, NJ.

Cristino, A. S., F. M. F. Nunes, C. H. Lobo, M. M. G. Bitondi, Z. L. P. Simoes, L. da Fontoura Costa, H. M. G. Lattorff, R. F. A. Moritz, J. D. Evans, and K. Hartfelder. 2006. Caste development and reproduction: a genome-wide analysis of hallmarks of insect eusociality. Insect Molecular Biology 15: 703-714.

Criswell, J., K. Pinkston, J. Igleheart, and S. Wells. 2001. Choosing a termite control service. OSU Extension Fact Sheet F-7308. Oklahoma State University, Stillwater, OK. 4 Pp.

Danty, E., G. Arnold, T. Burmester, J.-C. Huet, D. Huet, J.-C. Pernollet, and C. Masson. 1998. Identification and developmental profiles of hexamerins in antenna and hemolymph of the honeybee, *Apis mellifera*. Insect Biochemistry and Molecular Biology 28: 387-397.

Domon, B., and R. Aebersold. 2006. Challenges and opportunities in proteomics data analysis. Molecular and Cellular Proteomics 5: 1921-1926.

Donnelly, B. E. 2003. Proteomic analysis of aphid-wheat interactions. Ph.D. Dissertation. Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK. 196 Pp.

Donnelly, B. E., R. D. Madden, P. Ayoubi, D. R. Porter, and J. W. Dillwith. 2005. The wheat (*Triticum aestivum* L.) leaf proteome. Proteomics 5: 1624-1633.

DowAgroSciences. 1998-2004. Is my home at risk? <http://www.dowagro.com/sentricon/us/risk/index.htm>. Accessed: 18 February 2004

Dwek, M. V., and A. A. Alaiya. 2003. Proteome analysis enables separate clustering of normal breast, benign breast and breast cancer issues. British Journal of Cancer 89: 305-307.

ESA. 2004. Common names of insects and related organisms. Entomological Society of America. Lanham, MD.

Ferro-Luzzi Ames, G., and K. Nikaido. 1976. Two-dimensional gel electrophoresis of membrane proteins. Biochemistry 15: 616-623.

Fichmann, J. 1999. Advantages of immobilized pH gradients, pp. 173-188. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, NJ.

Finkelstein, A. V., and O. B. Ptitsyn. 2002. Protein Physics: A Course of Lectures. Academic Press, Boston. 353 Pp.

Gene Ontology Consortium. 1999-2007. Gene ontology. <http://www.geneontology.org/>. Accessed: 28 March 2007

Giavalisco, P., E. Nordhoff, T. Kreitler, K.-D. Kloppel, H. Lehrach, J. Klose, and J. Gobom. 2005. Proteome analysis of *Arabidopsis thaliana* by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry. Proteomics 5: 1902-1913.

Gilbert, L. I., N. A. Granger, and R. M. Roe. 2000. The juvenile hormones: historical facts and speculations on future research directions. Insect Biochemistry and Molecular Biology 30: 617-644.

Giorgianni, F., D. M. Desiderio, and S. Beranova-Giorgianni. 2003. Proteome analysis using isoelectric focusing in immobilized pH gradient gels followed by mass spectrometry. Electrophoresis 24: 253-259.

Gold, R. E., H. N. Howell, Jr., and G. J. Glenn. 1999. Subterranean termites. House and Landscape Pests B-6080. Texas A&M University, College Station, TX. 8 Pp.

Gorg, A., and W. Weiss. 1999. Analytical IPG-Dalt, pp. 189-195. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, NJ.

Gras, R., P. Hernandez, M. Muller, and R. D. Appel. 2003. Scoring functions for mass spectrometric protein identification, pp. 477-485. *In*: P. M. Conn [ed.], Handbook of Proteomic Methods. Humana Press, Inc., Totowa, NJ.

Graves, P. R., and T. A. J. Haystead. 2003. Proteomics and the molecular biologist, pp. 3-16. *In*: P. M. Conn [ed.], Handbook of Proteomic Methods. Humana Press, Inc., Totowa, NJ.

Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold. 1999a. Correlation between protein and mRNA abundance in yeast. Mollecular and Cellular Biology 19: 1720-1730.

Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold. 1999b. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. Nature Biotech 17: 994-999.

Hartson, S., T. Prince, P. Ayoubi, J. Rogers, J. Shao, U. Melcher, S. Hudiburg, A. Williams, and S. White. 2003. MALDI-TOF Mass Spectrometry and Introduction to Proteomics. Oklahoma State University, Stillwater, OK. 179 Pp.

Hernandez, R., C. Nombela, R. Diez-Orejas, and C. Gil. 2004. Two-dimensional reference map of *Candida albicans* hyphal forms. Proteomics 4: 374-382.

Hoefer Scientific Instruments. 1994. Protein Electrophoresis: Applications Guide. San Francisco, CA. 106 Pp.

Honigberg, B. M. 1970. Protozoa associated with termites and their role in digestion, pp. 1-36. *In*: K. Krishna and F. M. Weesner [eds.], Biology of Termites. Academic Press, New York.

Hoy, M. A. 2003. Insect Molecular Genetics: An Introduction to Principles and Applications. Academic Press, New York. 544 Pp.

Hu, Y., G. Wang, C. G. Y. J., X. Fu, and S. Yao. 2003. Proteome analysis of *Saccharomyces cerevisiae* under metal stress by two-dimensional gel electrophoresis. Electrophoresis 24: 1458-1470.

Immler, D., S. Greven, and P. Reinemer. 2006. Targeted proteomics in biomarker validation: detection and quantification of proteins using a multi-dimensional peptide separation strategy. Proteomics 6: 2947-2958.

Inoue, T., O. Kitade, T. Yoshimura, and I. Yamaoka. 2000. Symbiotic associations with protists, pp. 275-288. *In*: T. Abe, D. E. Bignell and M. Higashi [eds.], Termites: Evolution, Sociality, Symbioses, Ecology. Kluwer Academic Publishers, Boston, MA.

James, P., M. Quadroni, E. Carafoli, and G. Gonnet. 1993. Protein identification by mass profile fingerprinting. Biochemical and Biophysical Research Communications 195: 58-64.

Jensen, O. N., M. Wilm, A. Shevchenko, and M. Mann. 1999. 2-D proteome analysis protocols. Methods in Molecular Biology 112: 513-530.

Kabiri, K., H. Omidian, S. A. Hashemi, and M. J. Zohuriaan-Mehr. 2003. Synthesis of fast-swelling superabsorbent hydrogels: effect of crosslinker type and concentration on porosity and absorption rate. European Polymer Journal 39: 1341-1348.

Kahn, P. 1995. From genome to proteome: looking at a cell's proteins. Science 270: 369-370.

Kard, B. M., J. L. Etheridge, E. J. Mallette, and N. M. Rich. 2003. Procedures for preparing subterranean termites for laboratory studies (Isoptera: Rhinotermitidae). Sociobiology 41: 495-511.

Klose, J. 1999. Large-gel 2-D electrophoresis, pp. 147-172. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totawa, NJ.

Klose, J., and U. Kobalz. 1995. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis 16: 1034-1059.

Knowles, M. R., S. Cervino, H. A. Skynner, S. P. Hunt, C. de Felipe, K. Salim, Meneses-Lorente, G. McAllister, and P. C. Guest. 2003. Multiplex proteomic analysis by two-dimensional differential in-gel electrophoresis. Proteomics 3: 1162-1171.

Kolin, A. 1955. Isoelectric spectra and mobility spectra: a new approach to electrophoretic separation. Proceedings of the National Academy of Sciences 41: 101-110.

Krishna, K. 1969. Introduction, pp. 1-17. *In*: K. Krishna and F. M. Weesner [eds.], Biology of Termites. Academic Press, New York.

Krishna, K. 1989. An introduction to the study of insects, pp. 875. *In*: D. J. Borror, C. A. Triplehorn and N. F. Johnson [eds.], 6th ed. Saunders College Publishing, Fort Worth, TX.

Kuster, B., and M. Mann. 1999. ^{18}O-Labeling of n-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. Analytical Chemistry 71: 1431-1440.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Laine, L. V., and D. J. Wright. 2003. The life cycle of *Reticulitermes* spp. (Isoptera: Rhinotermitidae): what do we know? Bulletin of Entomological Research 93: 267-278.

Lewis, D. K., D. Spurgeon, T. W. Sappington, and L. L. Keeley. 2002. A hexamerin protein, AgSP-1, is associated with diapause in the boll weevil. Journal of Insect Physiology 48: 887-901.

Liebler, D. C. 2002. Introduction to Proteomics: Tools for the New Biology. Humana Press, Inc., Totowa, NJ. 198 Pp.

Marekov, L. N., and P. M. Steinert. 2003. Charge derivatization by 4-sulfophenyl isothiocyanate enhances peptide sequencing by post-source decay matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Journal of Mass Spectrometry 38: 373-377.

Mathews, C. K., K. E. van Holde, and K. G. Ahern. 2000. Biochemistry. Addison Wesley Longman, New York. 1,186 Pp.

Matrix Science Ltd. 2007. <http://www.matrixscience.com/>. Accessed: 10 March 2007

Matsui, N. M., D. M. Smith-Beckerman, and L. B. Epstein. 1999. Staining of preparative 2-D gels: Coomassie blue and imidazole-zinc negative staining, pp. 307-311. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, N. J.

Miller, E. M. 1969. Caste differentiation in the lower termites, pp. 283-310. *In*: K. Krishna and F. M. Weesner [eds.], Biology of Termites. Academic Press, New York.

Miura, T. 2001. Morphogenesis and gene expression in the soldier-caste differentiation of termites. Insectes Sociaux 48: 216-223.

Moreira, C. K., M. d. L. Capurro, M. Walter, E. Pavlova, H. Biessmann, A. A. James, A. G. deBianchi, and O. Marinotti. 2004. Primary characterization and basal promoter activity of two hexamerin genes of *Musca domestica*. Journal of Insect Science 4: 1-10.

Musante, L., G. Caniano, and G. M. Ghiggeri. 1998. Resolution of fibronectin and other uncharacterized proteins by two-dimensional polyacrylamide electrophoresis with thiourea. Journal of Chromatography B 705: 351-356.

NanoDrop Technologies. 2006. <http://www.nanodrop.com/products.html>. Accessed: 11 October 2006

National Center for Biotechnology Information. 2006. NCBI Blast. <http://www.ncbi.nlm.nih.gov/BLAST/>. Accessed: 16 October 2006

Nielsen, P. A., J. V. Olsen, A. V. Podtelejnikov, J. R. Andersen, M. Mann, and J. R. Wisniewski. 2005. Proteomic mapping of brain plasma membrane proteins. Molecular and Cellular Proteomics 4: 402-408.

Nutting, W. L. 1990. Insecta: Isoptera, pp. 997-1032. *In*: D. L. Dindal [ed.], Soil Biology Guide. John Wiley & Sons, New York, New York.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. The Journal of Biological Chemistry 250: 4007-4021.

OSU Entomology and Plant Pathology. 2006. Entomology and Plant Pathology - OSU. <http://www.ento.okstate.edu/profiles/dillwith.htm>. Accessed: 15 March 2007

Palzkill, T. 2002. Proteomics. Kluwer Academic Publishers, Norwell, Massachusetts. 127 Pp.

Pandey, A., and M. Mann. 2000. Proteomics to study genes and genomes. Nature 405: 837-846.

Pappin, D. J. C., P. Hojrup, and A. J. Bleasby. 1993. Rapid identification of proteins by peptide-mass fingerprinting. Current Biology 3: 327-332.

Pennington, S. R., M. R. Wilkins, D. F. Hochstrasser, and M. J. Dunn. 1997. Proteome analysis: from protein characterization to biological function. Trends in Cell Biology 7: 168-173.

Perkins, D. N., D. J. C. Pappin , D. M. Creasy, and J. S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20: 3551-3567.

Petricoin, E. F., and L. A. Liotta. 2004. Proteomic approaches in cancer risk and response assessment. Trends in Molecular Medicine 10: 59-64.

Petsko, G. A., and D. Ringe. 2004. Protein Structure and Function. New Science Press, Ltd., Singapore. 195 Pp.

Posch, A., B. M. van den Berg, H. C. J. Burg, and A. Gorg. 1995. Genetic variability of carrot seed proteins analyzed by one- and two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 16: 1312-1316.

Rabilloud, T. 1999. Silver staining of 2-D electrophoresis gels, pp. 297-306. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, N. J.

Ramagli, L. S. 1999. Quantifying protein in 2-D PAGE solubilization buffers, pp. 99-103. *In*: A. J. Link [ed.], Methods in Molecular Biology. Humana Press, Inc., Totowa, NJ.

Redei, G. P. 2003. Encyclopedic Dictionary of Genetics, Genomics, and Proteomics. Wiley & Sons, Inc., Hoboken, NJ. 1379 Pp.

Rodriguez-Ortega, M. J., B. E. Grosvik, A. Rodriguez-Ariza, A. Goksoyr, and J. Lopez-Barea. 2003. Changes in protein expression profiles in bivalve molluscs (*Chamaelea gallina*) exposed to four model environmental pollutants. Proteomics 3: 1535-1543.

Rosenfeld, J., J. Capdevielle, J. C. Guillemot, and P. Ferrara. 1992. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Analytical Biochemistry 203: 173-179.

Scarselli, R., E. Donadio, M. G. Giuffrida, D. Fortunato, A. Conti, E. Balestreri, R. Felicioli, M. Pinzauti, A. G. Sabatini, and A. Felicioli. 2005. Towards royal jelly protoeme. Proteomics 5: 769-776.

Scharf, M. E., D. Wu-Scharf, X. Zhou, B. R. Pittendrigh, and G. W. Bennett. 2005a. Gene expression profiles among immature and adult reproductive castes of the termite *Reticulitermes flavipes*. Insect Molecular Biology 14: 31-44.

Scharf, M. E., C. R. Ratliff, D. Wu-Scharf, X. Zhou, B. R. Pittendrigh, and G. W. Bennett. 2005b. Effects of juvenile hormone III on *Reticulitermes flavipes*: changes in hemolymph protein composition and gene expression. Insect Biochemistry and Molecular Biology 35: 207-215.

Scheller, K., B. Fischer, and H. Schenkel. 1990. Molecular properties, functions and developmentally regulated biosynthesis of arylphorin in *Calliphora vicina*., pp. 155-162. *In*: H. H. Hagedorn, J. G. Hildebrand, M. G. Kidwell and J. H. Law [eds.], Molecular Insect Science. Plenum Press, New York.

Serrano, S. M. T., J. D. Shannon, D. Wang, A. C. M. Camargo, and J. W. Fox. 2005. A multifaceted analysis of viperid snake venoms by two-dimensional gel electrophoresis: an approach to understanding venom proteomics. Proteomics 5: 501-510.

Service, R. F. 2001. High-speed biologists search for gold in proteins. Science 294: 2074-2077.

Sharma, R., R. Sharma, H. Noda, and S. Komatsu. 2004. Proteomic analysis of brown planthopper: application to the study of carbamate toxicity. Insect Biochemistry and Molecular Biology 34,: 425-432.

Shevchenko, A., M. Wilm, O. Vorm, and M. Mann. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Analytical Chemistry 68: 850-858.

Simpson, R. J. 2003. Proteins and Proteomics: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York. 926 Pp.

Snyder, T. E. 1925. The origin of the castes in termites. Biological Society of Washington 22: 57-67.

Snyder, T. E. 1954. Order Isoptera: The Termites of the United States and Canada. National Pest Control Association, New York, New York. 64 Pp.

Spear, P. J. 1970. Principles of termite control, pp. 577-604. *In*: K. Krishna and F. M. Weesner [eds.], Biology of Termites. Academic Press, New York.

Spengler, B. 2001. The basics of matrix-assisted laser desorption, ionisation time-of-flight mass spectrometry and post-source decay analysis, pp. 33-53. *In*: P. James [ed.], Proteome Research : Mass Spectrometry. Springer, New York.

Sprenger, R. R., D. Speijer, J. W. Back, C. G. De Koster, H. Pannekoek, and A. J. G. Horrevoets. 2004. Comparative proteomics of human endothelial cell caveolae and rafts using two-dimensional gel electrophoresis and mass spectrometry. Electrophoresis 25: 156-172.

Stadler, F., and D. Hales. 2002. Highly-resolving two-dimensional electrophoresis for the study of insect proteins. Proteomics 2: 1347-1353.

Stewart, I. I., T. Thomson, D. Figeys, and H. S. Duewel. 2003. The use of ^{18}O Labeling as a tool for proteomic applications, pp. 145-179. *In*: P. M. Conn [ed.], Handbook of Proteomic Methods. Humana Press, Inc., Totowa, NJ.

Tarazka, J. A., R. Kurulugama, R. A. Sowell, S. J. Valentine, S. L. Koeniger, R. J. Arnold, D. F. Miller, T. C. Kaufman, and D. E. Clemmer. 2005. Mapping the proteome of *Drosophila melanogaster*: analysis of embryos and adult heads by LC-IMS-MS methods. Journal of Proteome Research 4: 1223-1237.

Telfer, W. H., and J. G. Kunkel. 1991. The function and evolution of insect storage hexamers. Annual Review of Entomology 36: 205-228.

Thorne, B. L., and J. F. A. Traniello. 2003. Comparative social biology of basal taxa of ants and termites. Annual Review of Entomology 48: 283-306.

UniProt Consortium. 2006. <http://www.pir.uniprot.org/>. Accessed: 28 March 2007

Veenstra, T. D. 2007. Global and targeted quantitative proteomics for biomarker discovery. Journal of Chromatography B 847: 3-11.

Veenstra, T. D., T. P. Conrads, B. L. Hood, A. M. Avellino, R. G. Ellenbogen, and R. S. Morrison. 2005. Biomarkers: mining the biofluid proteome. Molecular and Cellular Proteomics 4: 409-418.

Vierstraete, E., P. Verleyen, F. Sas, G. Van den Bergh, A. De Loof, L. Arckens, and L. Schoofs. 2004a. The instantly released *Drosophila* immune proteome is infection-specific. Biochemical and Biophysical Research Communications 317: 1052-1060.

Vierstraete, E., P. Verleyen, G. Baggerman, W. D'Hertog, G. Van den Bergh, L. Arckens, A. De Loof, and L. Schoofs. 2004b. A proteomic approach for the analysis of instantly released wound and immune proteins in *Drosophila melanogaster* hemolymph. Proceedings of the National Academy of Sciences 101: 470-475.

Vihinen, M. 2003. Bioinformatics in proteomics, pp. 419-428. *In*: P. M. Conn [ed.], Handbook of Proteomic Methods, 1st ed. Humana Press, Inc., Totowa, New Jersey.

Waldburg, N., T. Kahne, A. Reisenauer, C. Rocken, T. Welte, and F. Buhling. 2004. Clinical proteomics in lung diseases. Pathology - Research and Practice 200: 147-154.

Wang, Y. K., Z. Ma, D. F. Quinn, and E. W. Fu. 2001. Inverse ^{18}O labeling mass spectrometry for the rapid identification of marker/target proteins. Analytical Chemistry 73: 3742-3750.

Watson, B. S., V. S. Asirvatham, L. Wang, and L. W. Sumner. 2003. Mapping the proteome of barrel medic (*Medicago trunculata*). Plant Physiology 131: 1104 - 1123.

Weesner, F. M. 1970. Termites of the neartic region, pp. 477-525. *In*: K. Krishna and F. M. Weesner [eds.], Biology of Termites. Academic Press, New York.

Wilkins, M. R., E. Gasteiger, A. Bairoch, J.-C. Sanchez, K. L. Williams, R. D. Appel, and D. F. Hochstrasser. 1999. Protein identification and analysis tools in the ExPASy server, pp. 531-552. *In*: A. J. Link [ed.], 2-D Proteome Analysis Protocols. Humana Press, Inc., Totowa, NJ.

Yan, J. X., R. Wait, T. Berkelman, R. A. Harry, J. A. Westbrook, C. H. Wheeler, and M. J. Dunn. 2000. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis 21: 3666-3672.

Yan, Y., and G. Marriott. 2003. Analysis of protein interactions using fluorescence technologies. Current Opinion in Chemical Biology 7: 635-640.

Yao, X., A. Freas, J. Ramirez, P. A. Demirev, and C. Fenselau. 2001. Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. Analytical Chemistry 73: 2836-2842.

Yu, L.-R., T. P. Conrads, T. Uo, H. J. Issaq, R. S. Morrison, and T. D. Veenstra. 2004. Evaluation of the acid-cleavable isotope-coded affinity tag reagents: application to camptothecin-treated cortical neurons. Journal of Proteome Research 3: 469-477.

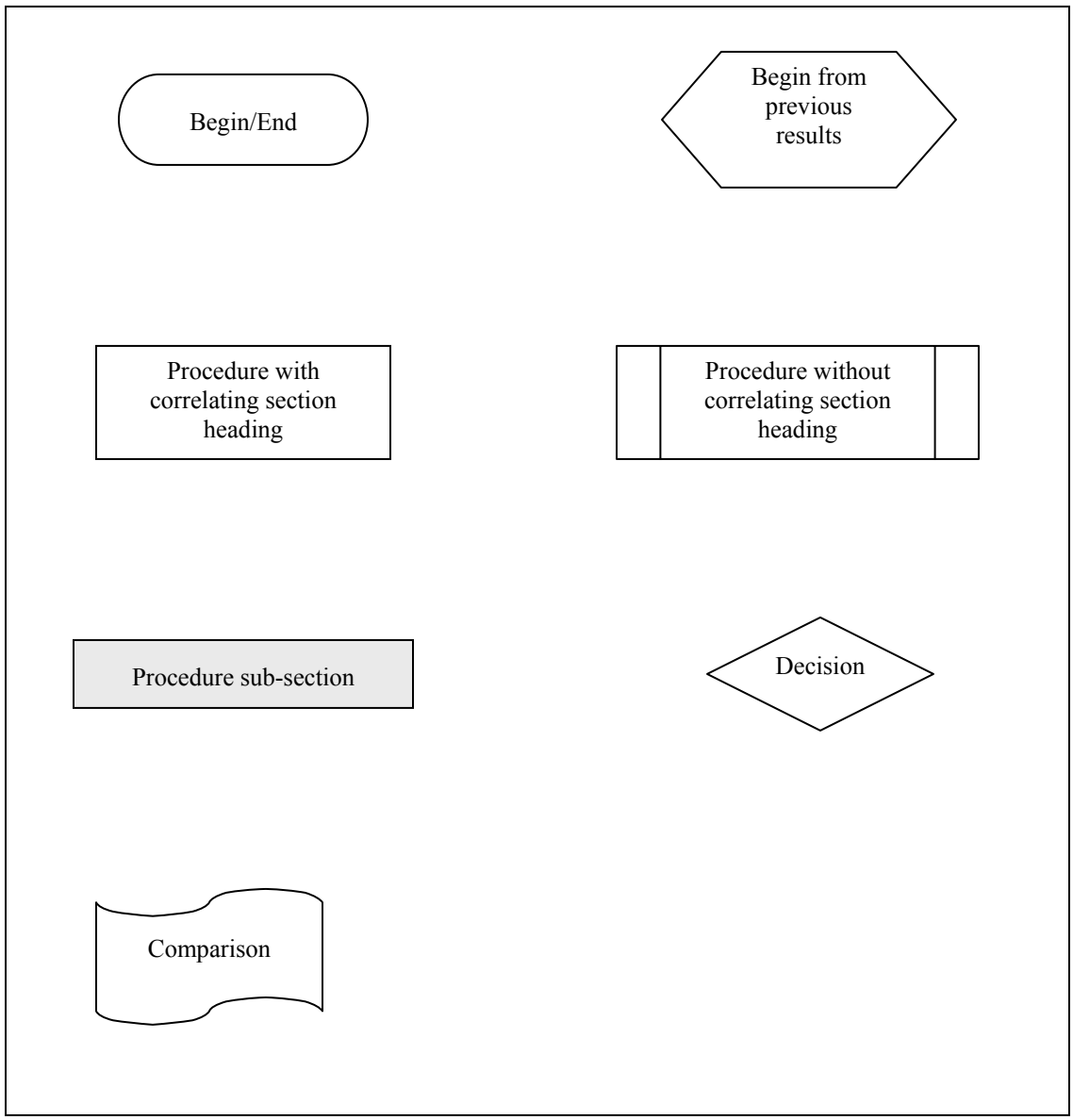
Zhou, X., F. M. Oi, and M. E. Scharf. 2006. Social exploitation of hexamerin: RNAi reveals a major caste-regulatory factor in termites. Proceedings of the National Academy of Sciences 103: 4499-4504.

Zivy, M., and D. de Vienne. 2000. Proteomics: a link between genomics, genetics, and physiology. Plant Molecular Biology 44: 575-580.

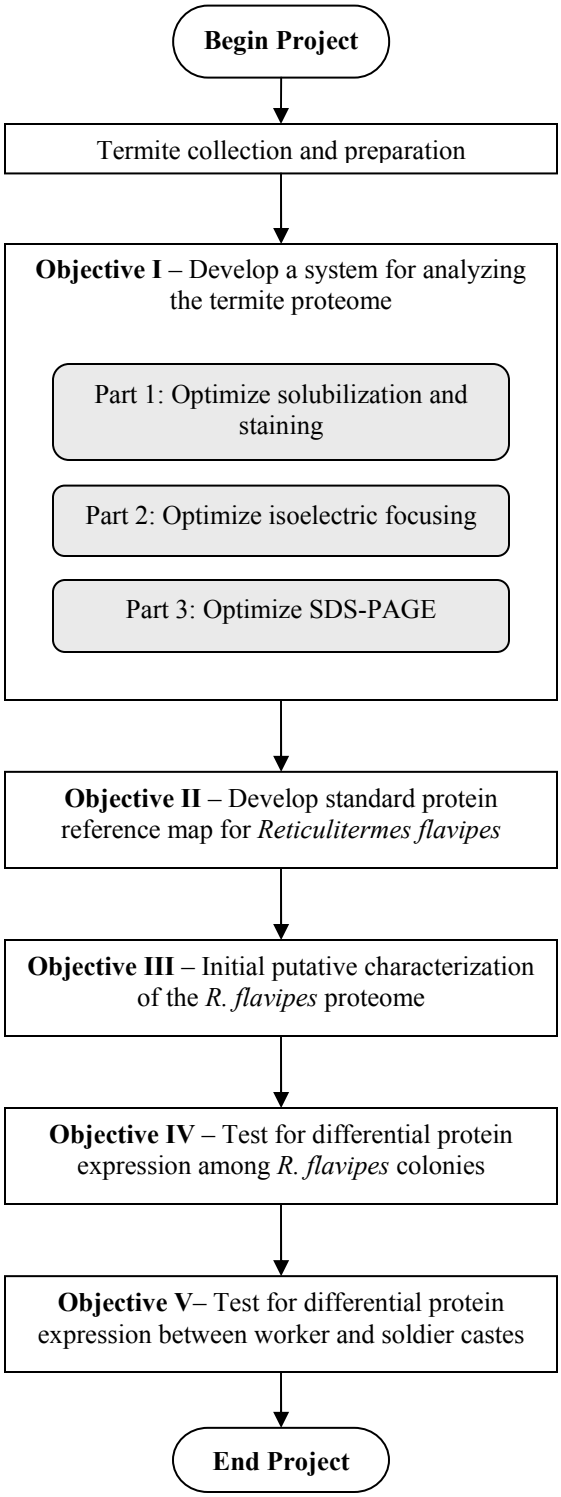
APPENDICES

Appendix A – Flowcharts

Legend

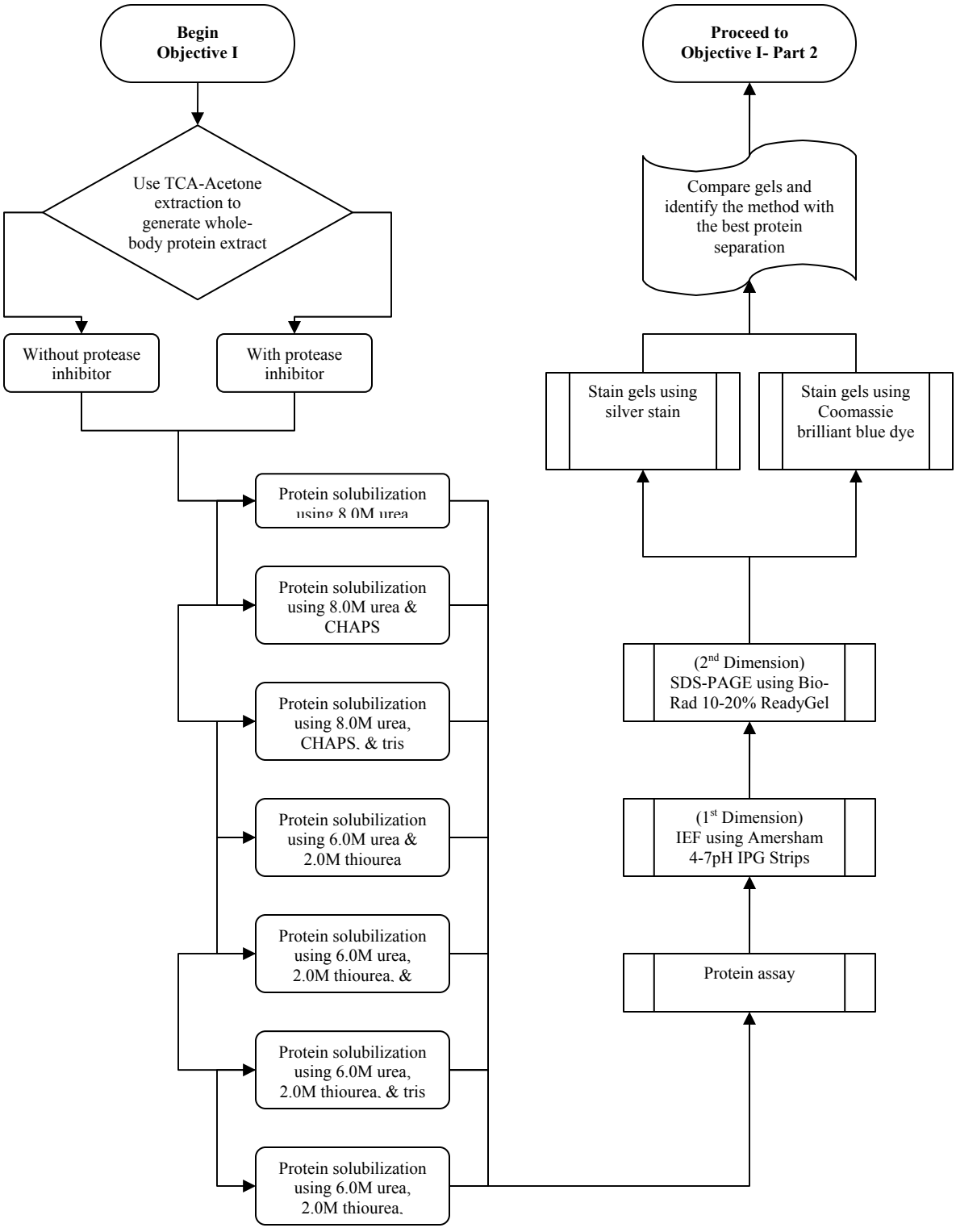


Flowchart Illustrating Overall Project



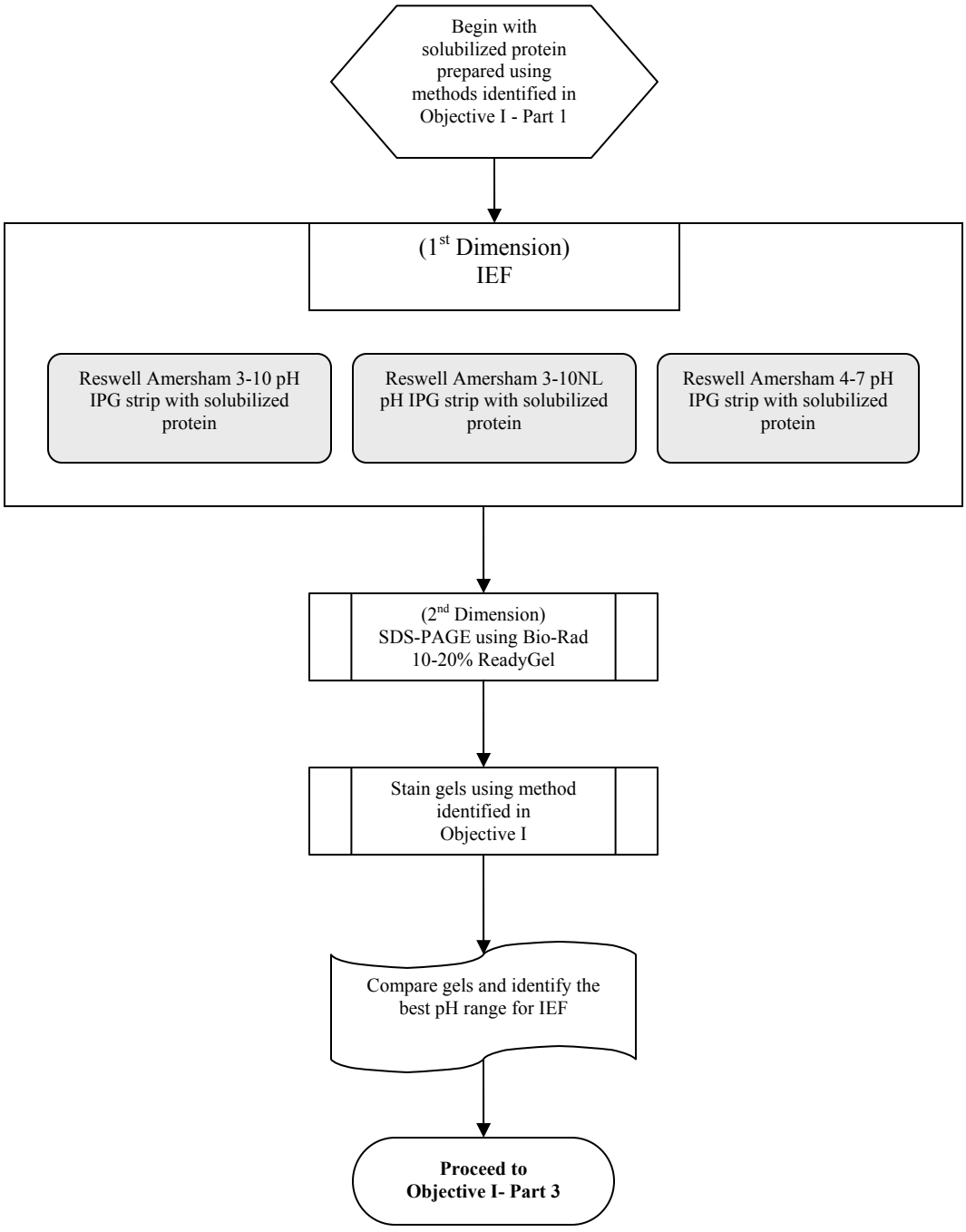
Flowchart Illustrating Objective I – Develop a system for analyzing the termite proteome

Part 1: Optimize solubilization and staining



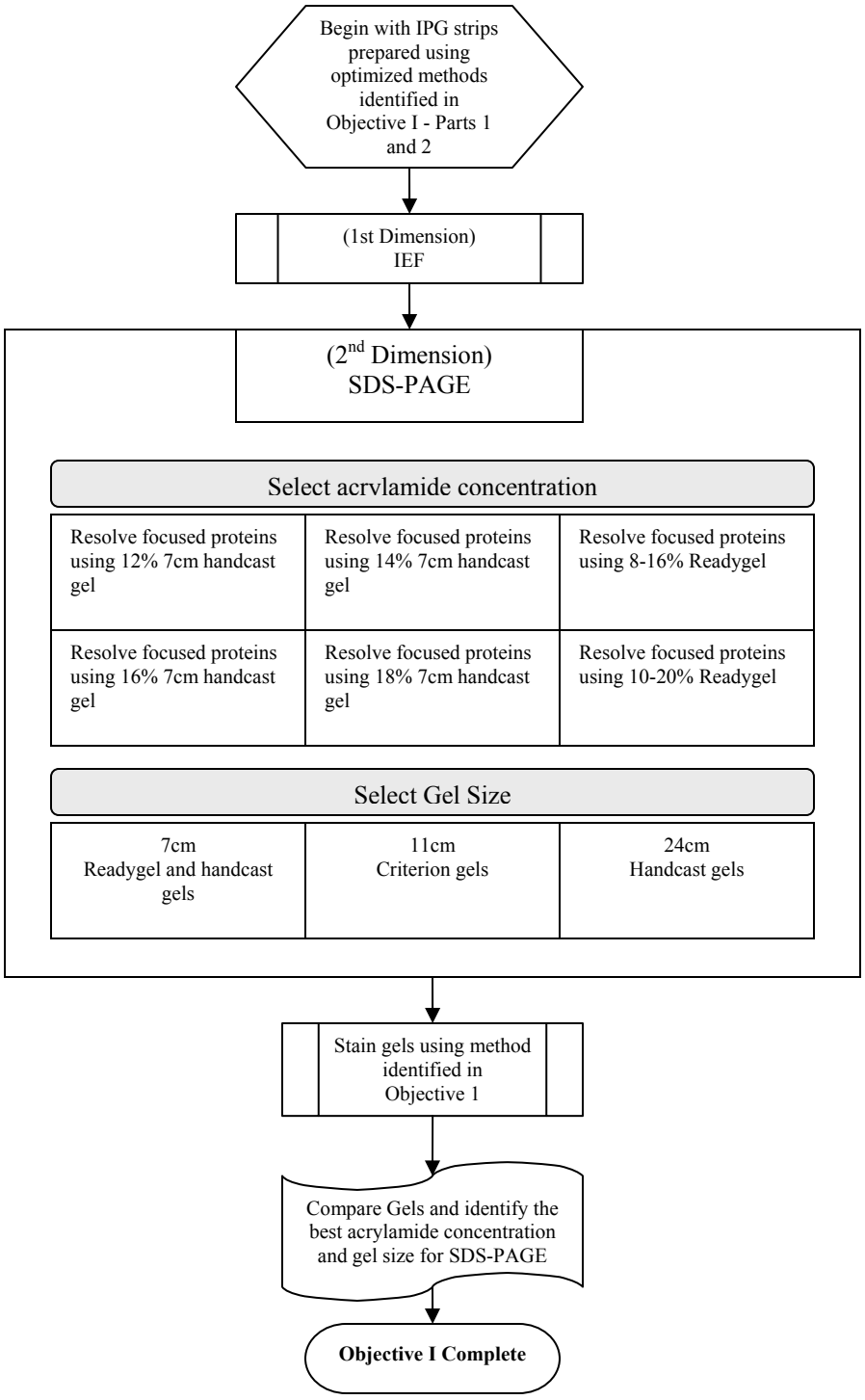
Flowchart Illustrating Objective I – Develop a system for analyzing the termite proteome

Part 2: Optimize IEF

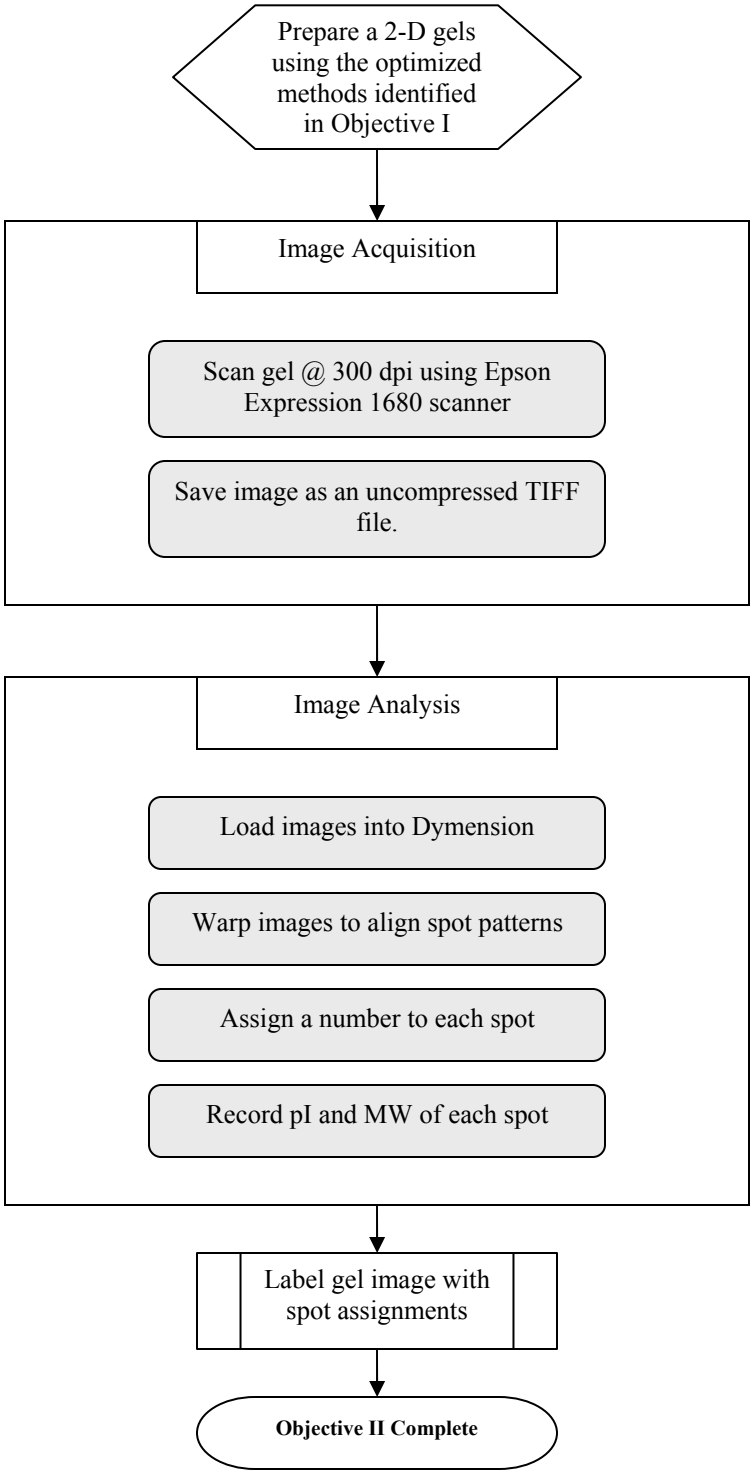


Flowchart Illustrating Objective I: Develop a system for analyzing the termite proteome

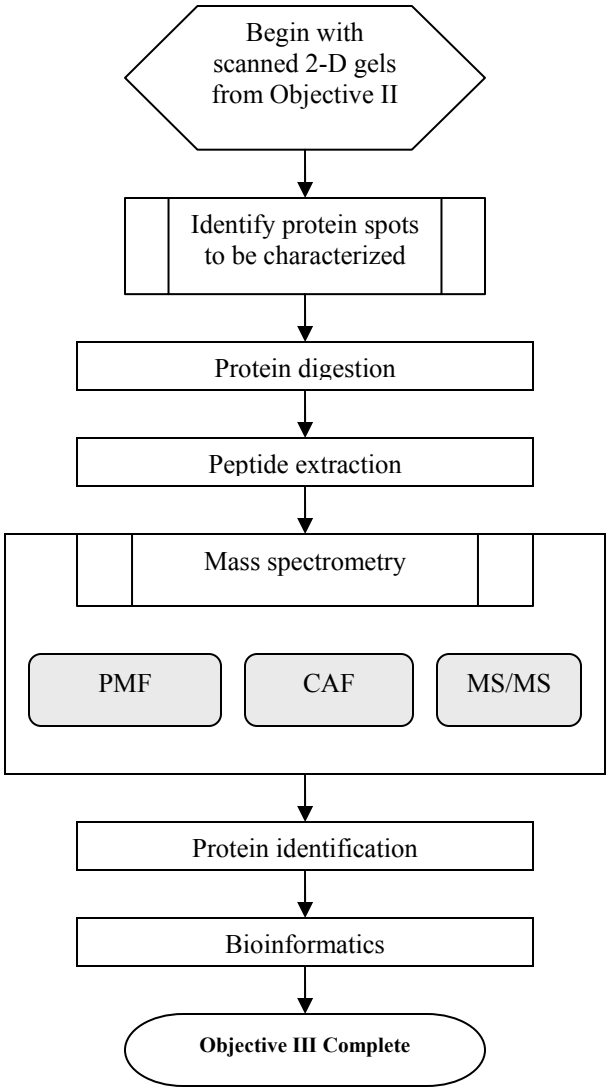
Part 3: Optimize SDS-PAGE



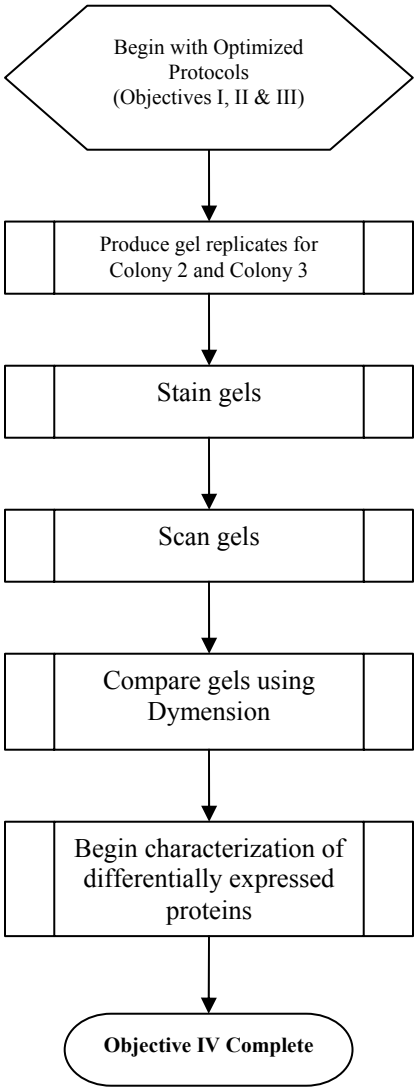
Flowchart Illustrating Objective II – Develop standard protein reference map for *Reticulitermes flavipes*



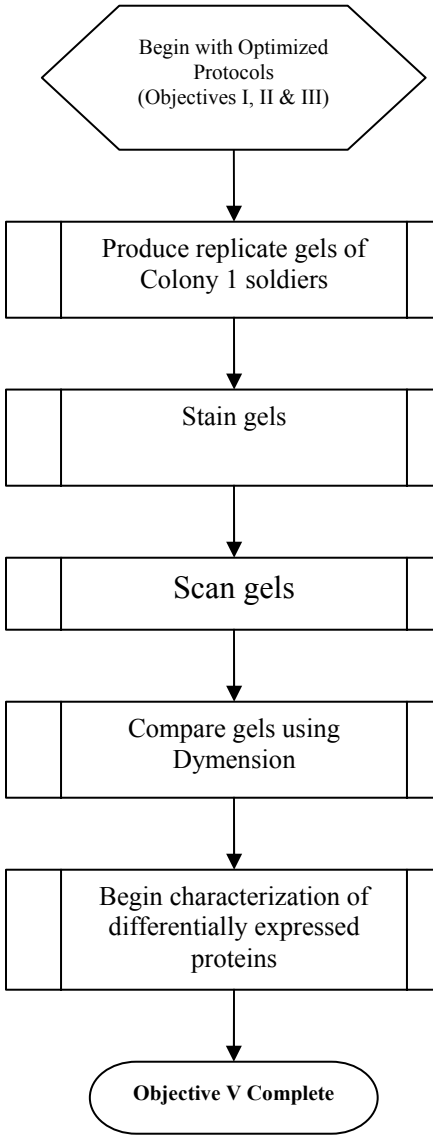
Flowchart Illustrating Objective III – Initial characterization of the putative *R. flavipes* proteome



Flowchart Illustrating Objective IV – Test for differential protein expression among *R. flavipes* colonies



Flowchart Illustrating Objective V –Test for differential protein expression between worker and soldier castes



Appendix B – Suggested Reagent List

Reagent		Chemical formula	Supplier	Unit	Part #
Acetic acid		$C_2H_4O_2$	Fisher	4L	A38-500
Acetone		C_3H_6O	Fisher	4L	A929-4
Acetonitrile (Optima)		C_2H_3N	Fisher	4L	A21-4
Acrylamide		C_3H_5NO	Amersham	1kg	17-1302-02
Agarose			Amersham	10g	17-0554-01
Ammonium bicarbonate		$CH_2O_3 \cdot H_3N$	Sigma	500g	A6141
Ammonium persulfate		$H_8N_2O_8S_2$	Amersham	25g	17-1311-01
B-mercaptoethanol		$C_2H_6O_5$	Bio-Rad		161-0710
BioRad protein assay dye reagant conc.			Bio-Rad	450mL	500-0006
Bis-acrylamide	(N,N^1-methylene bisacrylamide)	$C_7H_{10}N_2O_2$	Amersham	25g	17-1304-02
Brilliant blue (R-250)			Sigma	25g	B0149
Bromophenol blue			Sigma	1g	11,439-1
Butanol		$C_4H_{10}O$			
Calibration Mix			ABI		
CHAPS	(3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate)	$C_{32}H_{56}N_2Na_2O_{10}$	Amersham	1g	17-1314-01
DTT	(Dithiothreitol)	$C_4H_{10}O_2S_2$	Amersham	5g	17-1318-02
EDTA disodium salt	(Ethylenediaminetetraacetic acid disodium salt dehydrate)	$C_{10}H_{18}N_2Na_2O_{10}$	Amersham	100g	17-1324-01
Ethanol (200 proof)		C_2H_6O	Chem Store	4L	
Formalin (37%)		CH_2O	Sigma	500mL	F1268
Glycerol		$C_3H_8O_3$	Amersham	1L	17-1325-01
Glycine		$C_2H_5NO_2$	Bio-Rad	1kg	161-0718

Hydrochloric acid	(HCl)	ClH	Fisher	2.5L	A144-212
Immobiline^{TM} Drystrip cover fluid			Amersham	1L	17-1335-01
Iodoacetamide		C_2H_4INO	Sigma	5g	I1149
IPG buffer			Amersham	1mL	
Methanol		CH_4O	Fisher	20L	A434-20
Ovalbumin standard	(Albumin from hen egg white)		Sigma	Vial	O4757
Potassium ferricyanide		$K_3Fe(CN)_6$	Sigma	100g	P3667
Protease inhibitor cocktail			Sigma	1pkg	P2714
SDS	(Sodium dodecyl sulphate)	$C_{12}H_{25}NaO_4S$	Bio-Rad	100g	161-0301
Silver nitrate		$AgNO_3$	Sigma	100g	S0139
Sodium carbonate		Na_2CO_3	Sigma	1kg	S7795
Sodium thiosulfate		$Na_2S_2O_3 \cdot SH_2O$	Sigma	500g	S8503
TEMED	(N,N,N^1,N^1-tetra-methyl-ethylenediamine)	$C_6H_{16}N_2$	Bio-Rad	5mL	161-0800
Thiourea		CH_4N_2S	Sigma	50g	T8656
Trifluoroacetic acid		$C_2HF_3O_2$	Sigma	100g	T-1647
Trichloroacetic acid		$C_2HCl_3O_2$	Sigma	500g	T9159
Tris	(Tris(hydroxymethyl)aminomethane)	$C_4H_{11}NO_3$	Bio-Rad	1kg	161-0719
Triton X-100		$C_{16}H_{26}O_2$	Amersham	500mL	17-1315-01
Trypsin	(Sequencing Grade; modified)		Promega	100µg	V5113
Urea		CH_4N_2O	Amersham	500g	17-1319-01

End of Table

Appendix C – Links

Encyclopedia of Life
http://eol.sdsc.edu/

ExPASy Proteomics Server
http://au.expasy.org/

Folding @ Home: distributed computing
http://folding.stanford.edu/

Gene Ontology
http://geneontology.org

Harvard Institute of Proteomics
http://www.hip.harvard.edu/

Molecular Visualization Freeware: Protein Explorer, Chime, & RasMol
http://www.umass.edu/microbio/rasmol/

Mascot
http://www.matrixscience.com/

National Center for Biotechnology Information
http://www.ncbi.nlm.nih.gov/

NCBI Entrez Protein
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein

Online Analysis Tools
http://molbiol-tools.ca/

The Protein Database
http://www.rcsb.org/pdb/

Protein Explorer
http://www.umass.edu/microbio/chime/pe_beta/pe/protexpl/

Protein Information Resource
http://www-nbrf.georgetown.edu/

Protein Prospector
http://prospector.ucsf.edu/

PROWL
http://prowl.rockefeller.edu/

RSCB Protein Databank
http://www.rcsb.org/pdb/home/home.do;jsessionid=8DA087882A4B1F46027D522DE68DB154

UniProt Database
http://pir.uniprot.org

VITA

Charles Jerry Bowen, Jr.

Candidate for the degree of

Doctor of Philosophy

Thesis: *RETICULITERMES FLAVIPES* (ISOPTERA: RHINOTERMITIDAE) PROTEOMICS

Major Field: Entomology

Biographical:

Education: Received Bachelor of Science degree in Mathematics and a Master of Science degree in Biology from Midwestern State University, Wichita Falls, Texas in May 1994 and August 2002, respectively. Completed the requirements for the Doctor of Philosophy degree in Entomology from Oklahoma State University in May 2007.

Related Experience: Employed by Top-O-Texas Pest Control in Wichita Falls, Texas as a part-time employee from 1979 to 1983; as a pest control technician from 1983 to1990. Employed as a teaching assistant at Midwestern State University from 2000 to 2002 and at Oklahoma State University from 2002 to 2003. Employed as a research assistant at Midwestern State University from 2001 to 2002 and at Oklahoma State University from 2003 to present.

Professional Memberships: AAAS, American Arachnological Society, American Entomological Society, Central State Entomological Society, Entomological Society of America, North Texas Pest Control Association, Sanborn Entomology Club, Sigma Xi, Society of Southwestern Entomologist, and Texas Academy of Science.

Name: Charles Jerry Bowen, Jr. Date of Degree: May 2007

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: *RETICULITERMES FLAVIPES* (ISOPTERA: RHINOTERMITIDAE) PROTEOMICS

Pages in Study: 215 Candidate for the Degree of Doctor of Philosophy

Major Field: Entomology

Scope and Method of Study: The objectives of this study were to establish a system for analyzing the termite proteome, develop a standard protein reference map for *Reticulitermes flavipes*, begin putative characterization of the proteome, and compare differences among colonies and between the worker and soldier castes. The system was established using two-dimensional gel electrophoresis for protein separation and coomassie brilliant blue stain for protein visualization. Proteins were analyzed with MALDI-TOF mass spectrometry and tandem mass spectrometry (MS/MS). Resulting data were searched against protein databases for putative identification. A standard protein reference map was generated and used to compare differences among colonies. A reference map for the soldier caste was generated to allow comparison between worker and soldier castes.

Findings and Conclusions: Two-dimensional *Reticulitermes flavipes* protein maps were comparable demonstrating less than 10% of differentially expressed proteins among castes. However, comparison of two-dimensional protein maps of worker vs soldier castes demonstrated substantial quantities (~ 50%) of differentially expressed proteins. Mass spectrometry was undertaken on 310 worker proteins and an additional 142 differentially expressed soldier proteins. Overall, database searches yielded high-confidence putative identifications for 18.7% of the proteins. Comparison between MALDI-TOF and MS/MS yields revealed a five-fold increase in identifications from 7% to 34%, respectively. Hexamerins were the protein group most commonly identified from a termite protein source.

Advisor's Approval: Bradford M. Kard