7-2-2015

Employing Limited Next Generation Sequence Data for the Development of Genetic Loci of Phylogenetic and Population Genetic Utility

Lauren Evenstone
leven001, leven001@fiu.edu

DOI: 10.25148/etd.FIDC000108

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

Part of the [Bioinformatics Commons](#), [Computational Biology Commons](#), [Entomology Commons](#), [Genetics Commons](#), [Genomics Commons](#), and the [Molecular Genetics Commons](#)

Recommended Citation

Evenstone, Lauren, "Employing Limited Next Generation Sequence Data for the Development of Genetic Loci of Phylogenetic and Population Genetic Utility" (2015). *FIU Electronic Theses and Dissertations*. 2191.
<https://digitalcommons.fiu.edu/etd/2191>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

EMPLOYING LIMITED NEXT GENERATION SEQUENCE DATA FOR THE
DEVELOPMENT OF GENETIC LOCI OF PHYLOGENETIC AND POPULATION
GENETIC UTILITY

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

FORENSIC SCIENCE

by

Lauren A. Evenstone

2015

To: Dean Michael R. Heithaus
College of Arts and Sciences

This thesis, written by Lauren A. Evenstone, and entitled Employing Limited Next Generation Sequence Data for the Development of Genetic Loci of Phylogenetic and Population Genetic Utility, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Martin Tracey

Bruce McCord

Jeffrey Wells, Major Professor

Date of Defense: June 1, 2015

The thesis of Lauren A. Evenstone is approved.

Dean Michael R. Heithaus
College of Arts and Sciences

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2015

DEDICATION

I dedicate this work to my family, especially my partner, Eric Koschnick; my parents, Devon and Shareen Evenstone; and my sisters, Amanda and Allison. There are no words to express how grateful I am for all of your support and encouragement.

ACKNOWLEDGMENTS

I would like to thank Dr. Wells, my advisor, as well as Dr. Tracey and Dr. McCord, my committee members for their guidance through this project. It would not have been successful without their input. I would also like to thank the other students in my research laboratory. Fanchen Bao and Josh Smith contributed their time, their patience, and some of their flies to my project. I could not have accomplished this research without them. Finally, I would like to thank Anne Andere and Dr. Picard's lab at the University of Indiana-Purdue University Indianapolis for spending their Spring Break teaching me about bioinformatics and the software used in this project.

ABSTRACT OF THE THESIS

EMPLOYING LIMITED NEXT GENERATION SEQUENCE DATA FOR THE
DEVELOPMENT OF GENETIC LOCI OF PHYLOGENETIC AND POPULATION
GENETIC UTILITY

by

Lauren A. Evenstone

Florida International University, 2015

Miami, Florida

Professor Jeffrey Wells, Major Professor

Massively parallel high throughput sequencers are transforming the scientific research by reducing the cost and time necessary to sequence entire genomes. The goal of this project is to produce preliminary genome assemblies of calliphorid flies using Life Technologies' Ion Torrent sequencing and Illumina's MiSeq sequencing. I located, assembled, and annotated a novel mitochondrial genome for one such fly, the little studied *Chrysomya pacifica* that is central to one hypothesis about blow fly evolution. With sequencing data from *Chrysomya megacephala*, its forensically relevant sister species, much insight can be gained by alignments, sequence and protein analysis, and many more tools within the CLC Genomics Workbench software program. I present these analyses here of these recently diverged species.

TABLE OF CONTENTS

CHAPTER	PAGE
1. INTRODUCTION	1
1.1 SEQUENCING	1
1.1.1 SANGER SEQUENCING	1
1.1.2 ION TORRENT PGM	4
1.1.3 ILLUMINA MISEQ	10
1.2 CLC GENOMICS WORKBENCH	13
1.3 CHRYSOMYA	16
2 GENOMIC DNA PREPARATION AND ASSEMBLY	18
2.1 ION TORRENT PGM	18
2.2 ILLUMINA MISEQ	21
2.3 CLC GENOMICS WORKBENCH ASSEMBLY	22
3 MICROSATELLITES	24
3.1 MSATCOMMANDER	24
3.2 CLC GENOMICS WORKBENCH	27
3.3 PCR DEVELOPMENT	30
4 MITOCHONDRIAL GENOME	36
4.1 ASSEMBLY	36
4.2 CLC GENOMICS WORKBENCH ANALYSIS	38
4.3 OPEN SOURCE ANALYSIS	49
4.4 PHYLOGENETIC ANALYSIS	55
5 CONCLUSION	62
6 REFERENCES	64
7 APPENDICES	73

TABLE LIST

TABLE	PAGE
2.1.1: OUTPUT INFORMATION ON SUCCESSFUL CMOS WELL USAGE	20
2.3.1: ASSEMBLY SUMMARY REPORT METRICS	24
3.2.1: PRIMER SETS WITH OBSERVED AND EXPECTED SIZE RANGES	29
3.3.1: SUGGESTED CONCENTRATION AND VOLUMES FOR PCR MASTER MIX	30
3.3.2: PCR MASTER MIX RECIPES	31
3.3.3: THERMOCYCLER SPECIFICATIONS	31
3.3.4: POSITIONING OF THE SINGLEPLEX REACTIONS WITHIN THE THERMOCYCLER	32
3.3.5: RESULTS OF SINGLEPLEX PCR BY AGAROSE GEL ELECTROPHORESIS	34
4.1.1: CODING SEQUENCES WITHIN THE CHRYSOMYA PACIFICA MITOCHONDRIAL GENOME	37
4.2.1: COMPARISON OF BASE DIFFERENCES BETWEEN C. PACIFICA AND TWO C. MEGACEPHALA REFERENCE GENOMES	41
4.2.2: PROTEIN DOMAINS FOR ND1, ND4, ND5, CYTB, COX1	50
4.3.1: COMPARISON OF CODING REGION LENGTHS IN MITOS, PUBLISHED C. MEGACEPHALA GENOMES, AND C. PACIFICA	52
4.3.2: COMPARISON OF THE POSITIONS AND LENGTHS OF C. PACIFICA CODING REGIONS BETWEEN MITOS AND CLC	53
4.4.1: SOURCES OF SPECIES AND SPECIMENS USED IN PHYLOGENETIC ANALYSES	57

FIGURE LIST

FIGURE	PAGE
1.1.1.1: EXAMPLES OF ELECTROPHORETIC SEPARATION OF SANGER SEQUENCING PRODUCTS	3
1.1.1.2: COST OF GENOME SEQUENCING	4
1.1.2.1: WORKFLOW FOR ION TORRENT LIBRARY PREPARATION	5
1.1.2.2: COMPLEMENTARY METAL OXIDE SEMICONDUCTOR	8
1.1.3.1: SCHEMATIC OF LIBRARY PREPARATION AND BRIDGE AMPLIFICATION	11
1.1.3.2: SCHEMATIC OF REVERSE TERMINATOR CHEMISTRY AND TECHNOLOGY	12
1.2.1: A BUBBLE CAUSED BY A HETEROZYGOUS SNP OR A SEQUENCING ERROR	14
1.2.2: ILLUSTRATION OF PAIRED END READS AND SINGLE END READS	15
3.1.1: MSATCOMMANDER INTERFACE	25
3.1.2: EXAMPLE OF MSATCOMMANDER PRIMER 3 OUTPUT	26
4.2.1: HYDROPHOBICITY OVERLAYS OF C. PACIFICA, MEGA1, AND MEGA2	45
4.3.1: QUALITY SCORES OF MITOS DESCRIBED CODING REGIONS	54
3.4.1: PAIRWISE COMPARISON OF 13 CHRYSOMYA FLIES, LUCILIA SERICATA, AND COCHLIOMYIA HOMNIVORAX USING THE COX1 GENE FOR DIFFERENCES AND DISTANCE	59
6.4.2: PAIRWISE COMPARISON OF 13 CHRYSOMYA FLIES, LUCILIA SERICATA, AND COCHLIOMYIA HOMNIVORAX USING THE MITOCHONDRIAL GENOME FOR DIFFERENCES AND DISTANCE	59

FIGURE LIST

FIGURE	PAGE
6.4.3: PAIRWISE COMPARISON OF 13 CHRYSOMYA FLIES, LUCILIA SERICATA, AND COCHLIOMYIA HOMNIVORAX USING THE COX1 GENE FOR GAPS AND PERCENT IDENTITY	60
6.4.4: PAIRWISE COMPARISON OF 13 CHRYSOMYA FLIES, LUCILIA SERICATA, AND COCHLIOMYIA HOMNIVORAX USING THE MITOCHONDRIAL GENOME FOR GAPS AND PERCENT IDENTITY	60
6.4.5: PHYLOGENETIC TREES CREATED FROM A) COX1 AND B) WHOLE MITOCHONDRIAL GENOMIC SEQUENCE DATA	61

ABBREVIATIONS AND ACRONYMS

Adenine	A
Alanine	Ala
Arginine	Arg
Aspartamine	Asn
Aspartate	Asp
ATP Synthase F0 Subunit 6	ATP6
ATP Synthase F0 Subunit 8	ATP8
Base Pair	bp
Coding Domain Sequence	CDS
Complementary Metal-Oxide Semiconductor	CMOS
Cysteine	Cys
Cytochrome B Oxidase	CYTB
Cytochrome Oxidase Subunit 1	COX1
Cytochrome Oxidase Subunit 2	COX2
Cytochrome Oxidase Subunit 3	COX3
Cytosine	C
Degrees Celsius	°C
Deoxynucleoside Triphosphate	dNTP
Deoxyribonucleic Acid	DNA
Dideoxynucleoside Triphosphate	ddNTP
Gigabase	Gb
Glutamine	Gln
Glutamine	Glu

Glycine	Gly
Guanine	G
Histidine	His
Inter Simple Sequence Repeats	ISSR
Isoleucine	Ile
Large subunit rRNA	16s
Leucine	Leu
Lysine	Lys
Melting or Annealing Temperature	T _m
Messenger Ribonucleic Acid	mRNA
Methionine	Met
Mitochondrial DNA	mtDNA
NADH Dehydrogenase subunit 1	ND1
NADH Dehydrogenase subunit 2	ND2
NADH Dehydrogenase subunit 3	ND3
NADH Dehydrogenase subunit 4	ND4
NADH Dehydrogenase subunit 4L	ND4L
NADH Dehydrogenase subunit 5	ND5
NADH Dehydrogenase subunit 6	ND6
National Center for Biotechnology Information	NCBI
Next Generation Sequencing	NGS
Personal Genome Machine	PGM
Phenylalanine	Phe
Polymerase Chain Reaction	PCR

Post Mortem Interval	PMI
Primer Dimer	PD
Proline	Pro
Ribonucleic Acid	RNA
Ribosomal Ribonucleic Acid	rRNA
Serine	Ser
Short Tandem Repeats	STR
Small subunit rRNA	12s
Threonine	Thr
Thymine	T
Transfer Ribonucleic Acid	tRNA
Tryptophan	Trp
Tyrosine	Tyr
Valine	Val

1. INTRODUCTION

As the cost of DNA sequencing declines, it has become possible to apply a genomics approach to a wide variety of research questions. Next Generation Sequencing (NGS) platforms generate massive amounts of data much quicker and for far less cost than traditional Sanger Sequencing (Quail *et al.*; 2012, Bradnam *et al.*, 2013; Pettersson *et al.*, 2009). The sequencing data produced can vary in depth of coverage as well as in the totality of the genome represented (Sims, 2014). The sequencing depth can reflect aspects of genome structure such as nucleotide content and repetitive stretches but also in characteristic weaknesses in the sequencing technologies (Quail *et al.*, 2012).

1.1 SEQUENCING

1.1.1 SANGER SEQUENCING

The first large-scale genome reconstruction efforts used Sanger sequencing techniques. (Russell, 2006). By cutting up the genome into random, small fragments that are then amplified by cloning, a three billion base pair length genome like the human genome can be sequenced (www.sanger.ac.uk). At first, during what is often called manual sequencing, the sequencing reaction was carried out in four reaction tubes. Each tube contains DNA polymerase, a short primer, and the four deoxynucleoside triphosphates: deoxyadenosine triphosphate, deoxyguanosine triphosphate, deoxycytidine triphosphate, deoxythymidine triphosphate. The tubes differ in that in each one

type of the nucleotides exists in two forms. One is a radioactively labeled dideoxynucleoside triphosphate (ddNTP) at a lower concentration than the dNTPs. The tubes are placed in a thermocycler for sequencing. The primer anneals to the single stranded DNA fragment. The DNA polymerase adds complementary dNTPs to elongate a new DNA strand. Through cyclical changes in temperature, the initial DNA fragments are primed, elongated, and denatured back to a single stranded conformation, eventually creating millions of copies. With Sanger Sequencing, sometimes the incorporated base is a labeled ddNTP instead of dNTP. The labeled base causes elongation to stop. Depending on where in the sequence the ddNTP is introduced, the fragment will be that much shorter than if the whole fragment was amplified. These varied length fragments are then separated electrophoretically by size. With each reaction tube producing radioactively tagged fragments for each ddNTP, these can be run on polyacrylimide gels in four lanes (method from Anderson *et al.*, 2010). See the left side of Figure 1.1.1.1 for an example of a manual gel separation. A more recent version of this process uses fluorescently labeled ddNTPs, with a separate color for each base, so that only one sequencing reaction rather than four is used per sample. It is also much more common now to use PCR product rather than cloned DNA as the template for the sequencing reaction. Electrophoresis is now much more likely to be performed on a capillary rather than a gel-based instrument; this is an automated and more sensitive alternative to polyacrylamide gel electrophoresis. An example of Sanger Sequencing on capillary electrophoresis can be seen on the right side of Figure 1.1.1.1

FIGURE 1.1.1.1: EXAMPLES OF ELECTROPHORETIC SEPARATION OF SANGER SEQUENCING PRODUCTS

The left side shows a simulated polyacrylamide gel. The right side shows the same simulation as output from capillary electrophoresis.

(Image:http://upload.wikimedia.org/wikipedia/commons/3/3d/Radioactive_Fluorescent_Seq.jpg)

Sanger Limitations

The Sanger method is accurate but is time consuming and costly. The process is limited by the time necessary to clone these unknown fragments into vectors (Russell, 2006). Each fragment is on average 500-600 bases long. It took participants in the Human Genome Project years and millions of dollars to assemble one genome (Wetterstrand, 2014). In the graphs below, the years 2001 to 2007 refer to genome sequencing costs using Sanger Sequencing. The cost for supplies and labor in Sanger sequencing limited the field of genomics. Also, Sanger Sequencing is not sensitive enough for detection of clinically relevant low-level mutant somatic alleles or organisms (Anderson *et al.*, 2010).

New technologies offer speed and more extensive coverage than the traditional method, and are quickly replaced it in most genomic applications.

FIGURE 1.1.1.2: COST OF GENOME SEQUENCING

Cost (log) per 1 million bases and cost per genome during the Human Genome Project. Sanger Sequencing used until 2007. Next Generation Technologies used after 2007. (Wetterstrand, 2014)

1.1.2 ION TORRENT PGM

Ion Torrent Library Preparation

The Florida International University's DNA Core, run by Dr. Paul Sharp, provides sequencing services to campus and community research groups.

Although members of my laboratory extracted the four different specimens' genomic DNA, all steps in the process of sequencing were performed by Dr.

Sharp. All reagents mentioned are proprietary and their contents are not

available. A schematic of the workflow can be found in Figure 1.1.2.1.

FIGURE 1.1.2.1: WORKFLOW FOR ION TORRENT LIBRARY PREPARATION

(Image: lifetechnologies.com)

Ion Torrent is a voltage based sequencer. The extracted genomic DNA is enzymatically sheared using Life Technologies' Ion Shear Enzyme Mix. The reaction is contained and controlled by adding Ion Shear Buffer to halt the reaction. The fragments are then purified using Agencourt AMPure XP magnetic beads. The fragments adhere to the beads and are washed in ethanol to remove any potential contaminants such as the Ion Shear enzymes. The fragments are removed via the addition of a low TE buffer. The supernatant containing the

fragments was injected into an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) to determine the size after shearing and purification.

The ends are repaired and ligated with the two adapters, A and P1. The A adapter has a biotinylated motif that will be used to attach the fragment to the streptavidin beads in a later step. The P1 adapter will attach to the Ion Sphere Particles used in sequencing. An additional barcode sequence can be incorporated as well. This barcode would be of benefit to those wishing to sequence more than one individual on the same chip (chip to be discussed later). The barcodes would be used to sort the subsequent sequencing reads into each individual. Barcoding was not used in my samples as each specimen had its own chip. Also during the adapter ligation step, nick ends are prevented because both ends of the sequence are now composed of adapters.

The ligation step is followed by another purification step which is identical to the first purification step discussed above. The prepared library was run on an agarose gel with a size ladder. The target length of around 480bp (400 bp new sample sequence plus 80 bp adapters).was excised from the gel, and the agarose was removed.

The Ion Torrent One Touch ES is a robotic pipetting system that adds the proprietary reagents in a precise manner. This system ensures that no pipetting errors occur and also that extremely accurate amounts of the reagents are added to the solution. The prepared library is then transferred to the Ion Torrent One Touch 2. Here, the library will be introduced to the Ion Sphere Particles (ISPs)

that will be the solid substrate for sequencing. In theory, one sequence will attach to each ISP. As can be seen in Table 3.2.1, some beads end up with more than one sequence. These beads follow a serpentine path through a channel based emulsion PCR. Here, the location within the path will determine what temperature the ISPs are subjected to, and what aspect of a normal PCR cycle they are in: denaturation, annealing, and elongation. Following an initial five minutes at 95°C for denaturing, the cycle runs as such: denature 15 seconds at 95°C, anneal 15 seconds at 58°C, extend one minute at 70°C, and a final hold temperature of 4°C. In total, nine cycles occur where each ISP becomes covered in identical sequences. A third purification step is then performed and the fragment lengths are again checked on the Bioanalyzer.

Ion Torrent Sequencing

The solution of ISPs containing my library was inserted into a specialized syringe used to introduce the spheres to the microchip. Ion Torrent uses complementary metal-oxide semiconductor (CMOS) chips that contain millions of microscopic wells. Though many models are available, The DNA Core chose the 318 which contains 13 million wells. Within each well, an ISP is deposited. Each well acts as a separate sequencer and contains its own voltage sensor. The wells have small pores in the bottom, below the sensor, that serve to drain the various wash fluids. A detailed schematic of the CMOS chip can be found in Figure 1.1.2.2. The chip is inserted into the Ion Torrent Personal Genome Machine (PGM).

FIGURE 1.1.2.2: COMPLEMENTARY METAL OXIDE SEMICONDUCTOR

a). Image of the 318 Chip used (Image: LifeTechnologies.com). b). Schematic of the microfluidics and voltage monitoring system within a single well (Image: Rothberg *et al.*, 2011).

The Ion Torrent PGM is an automated system that uses microfluidics and minute voltage changes to record the incorporation of bases. Four vials containing one of each deoxynucleoside triphosphates: dGTP, dCTP, dTTP and dATP. A solution of dNTPs is washed over the chip. The incorporation of a base causes a decrease in the pH and an increase in the relative voltage following the release of a proton (Rothberg *et al.*, 2011). The voltage spike is read by the sensor beneath the well. If a base is incorporated in a well, its change in voltage is recorded for that specific well. The PGM cycles between dNTPs and a wash solution which removes remnant dNTPs that may still be present.

In total, 850 flows were used to sequence the four samples submitted to the DNA Core- each on its own chip. This many cycles allowed for the possible incorporation of 425 bases with intervening wash flows. Each well is represented

independently of all others providing a massively parallel sequencing system that takes only a couple hours (Quail *et al.*, 2012).

Output reads from each of the runs were analyzed by the Ion Torrent Software. Some reads were eliminated because of their lower PHRED score. The P represents the base-calling error probability. The log relationship creates the Q or PHRED score. The equation is available below, equation (1).

$$(1) \quad Q = -10 \log_{10} P$$

As an example, a score of 20 would indicate that there is a 1 in 100 chance that a base was called incorrectly, i.e. a 99% correct. In a survey of different sequencing platforms, Quail *et al.* (2012) found the average error rate for Ion Torrent PGM is 1.71%. The Ion Torrent software chose to eliminate 16.9-20.6% of the reads generated because of an unspecified quality threshold. The software also removed reads from wells that contained polyclonal ISPs which contain more than one initial sequence. In the end, 57.3-62.1% of the sequences were available. See Table 3.2.1 for the output values for each specimen. The final average read length of the raw data was around 250 and was much shorter than the anticipated 400 bases. The reads are single end reads. Other technologies offer paired end reads which uses a predetermined sequence length and sequences it from both ends. The imbedded metadata within the sequences provides insight into the distance between paired ends and offers more information on how these reads are to be assembled.

1.1.3 ILLUMINA MISEQ

Illumina MiSeq Library Preparation

All library preparation and sequencing was performed by the North Carolina State University Genomics Lab. Figure 1.1.3.1 shows the process of bridge amplification to create clusters of double-stranded DNA fragments.

The genomic DNA is fragmented and adaptors are attached to the ends: each end with a different adaptor. The fragments are washed over the flow cell which contains a lawn of oligonucleotides that complement the adaptors which causes the fragments to anchor at these points. During bridge amplification, the fragments bend down and anchor the other adaptor the as well. The system provides dNTPs to elongate a complementary sequence. The double stranded fragment is denatured, and the process repeats. This process eventually forms clusters of identical fragments that will later be sequenced.

FIGURE 1.1.3.1: SCHEMATIC OF LIBRARY PREPARATION AND BRIDGE AMPLIFICATION

(Image: illumina.com)

Illumina MiSeq Sequencing

The flow cell, with the attached clusters of amplified sequences, is washed with all four fluorescent dNTPs, primers, and enzymes. The primers adhere to the unattached end of the fragments. Depending on the sequence of the specific cluster, one of the four dNTPs will attach to the sequence adjacent to the newly bound primer. Using a laser to excite the dNTPs, each cluster will fluoresce one

of four colors, and the colors will be recorded by the imaging software. The fluorescent label is then removed from the base. All four dNTPs are again washed over the flow cell and the cycle repeats itself. See Figure 1.1.3.2

FIGURE 1.1.3.2: SCHEMATIC OF REVERSE TERMINATOR CHEMISTRY AND TECHNOLOGY

(Image: illumina.com)

Illumina Limitations

The Illumina sequencing platform can sequence a massive amount of DNA fragments at the same time. This technology, however, depends on the visualization of the different colonies of amplified reads. The need for a charged coupled device (CCD) camera can increase sequencing time (Rothberg *et al.*, 2011). The system must excite the fluorescent tags on the freshly incorporated bases, and then record what color is emitted at each of the thousands of colonies. The Ion Torrent system uses voltage fluctuations and CMOS chips that have preset well locations. The Illumina flow cell has a lawn of adapter ends that attach to the DNA fragments. The resulting clusters from the amplification process are randomly dispersed across the flow cell. This lack of organization may impede on efficiency of the CCD camera to locate and record each incorporation.

1.2 CLC GENOMICS WORKBENCH

Downstream analysis of NGS data can be hindered by the computational capacity necessary to assess the magnitude of data produced in massively parallel processing. Both platforms employed in the current project produce millions of reads with an average length of about 250 bases. These reads contain repetitive motifs and some sequencing errors. In the absence of a reference genome, the *De novo* assembly algorithm in CLC Genomics Workbench was employed.

De novo Assembly

The CLC Genomics Workbench uses a De Bruijn algorithm to assemble the short reads generated from NGS into long contiguous sequences called contigs. Within the analysis parameters, the word size and bubble size can be modified and optimized to each assembly. The word size refers to the minimum length of sequence that must overlap between reads. The word window shifts along the overlapped sequences. If overlapping reads diverge in sequence, caused by either sequencing errors or the presence of a heterozygous single nucleotide repeat (SNP), a bubble is created. See Figure 1.2.1. The program creates a path from the reads and these paths become the contigs. The second stage maps the same reads back to the newly formed contigs.

FIGURE 1.2.1: A BUBBLE CAUSED BY A HETEROZYGOUS SNP OR A SEQUENCING ERROR.

(Image: CLCbio.com)

Illumina MiSeq uses paired end technology. Paired end sequencing uses much longer initial fragments. By sequencing the first and last 250 base pairs and incorporating metadata into these reads, the sequencer produces more information spanning longer distances than simple single end reads (Quail et al., 2012). The assembly software assesses both paired end read sequences along with the information regarding the size of the gap that separates them and their

orientation with regards to one another. This is illustrated in Figure 1.2.2. The gap is represented as an arch connecting the two reads. When used in conjunction with the millions of other paired end reads produced in sequencing, large stretches of the genome can be organized. The three sections, A, B, and C, in Figure 1.2.2 show how only a few of these pairs (below) correlate to the reads necessary to construct the same sequence with single end reads. The green numbered sequence in Figure 1.2.2 is a schematic of regions in the genome.

The single end reads created using the Ion Torrent PGM platform offer less information in each read than Illumina's MiSeq paired end reads.

FIGURE 1.2.2: ILLUSTRATION OF PAIRED END READS AND SINGLE END READS

Assembling and annotating a nearly complete reference genome has often involved large scale collaboration between several research groups. For example the giant panda (*Ailuropoda melanoleura*) genome was sequenced,

assembled, and published via the cooperation of 26 laboratories (Li *et al.*, 2010). However NGS technology can also make it easier for one or a small group of scientists to undertake more limited and less costly projects than the large genome projects. In other words an incomplete assembly with relatively little coverage can be explored for genes and their regulatory units. One application of this approach is to target a locus with high copy number because it will show substantial coverage, and can therefore be easily identified and assembled, even with a small data set. One can also use existing sequence information, *e.g.*, from a closely related species, to search genome fragments for a locus of interest. I consider the possible utility of both approaches for advancing our knowledge of carrion flies in the genus *Chrysomya*.

1.3 *Chrysomya*

Chrysomya megacephala was once thought to exist as both a normal form and derived form. The normal form was found in dark, densely forested areas of Pacific Island nations. The derived form adapted to disturbed ecosystems that increase agricultural output for local peoples (Kurahashi, 1982). The derived form adapted to live not only in secondary forests and in the burned down shadows of former forest, but evolved to survive and thrive in urban settings alongside humans (Kurahashi, 1982). We now call the normal form *Chrysomya pacifica* and the derived form *Chrysomya megacephala* (Bharti & Kurahashi, 2009). While *C. pacifica* remains in the forested areas of islands to the east of Papua New

Guinea, the urbanized sister species *C. megacephala* has spread to Australia, mainland Asia, Africa, South America, and North America (Wells, 1991).

The two species likely diverged in New Guinea within the last 5,000 years (Kurahashi, 1982; Jeffrey Wells, unpublished data). Agriculture in this region began roughly 9,000 years ago with widespread deforestation practices by around 4,000 years ago (Powell, 1982). This time frame allows for more than 100,000 generations to occur within the *C. megacephala* population (Jeffrey Wells, unpublished data).

Chrysomya females usually deposit their eggs on carrion or less often on a live animal (Norris, 1965). Some species, including *C. megacephala*, are also highly attracted to both human waste and human food, and they are a vector of enteric diseases (Wells, 1991). Because the larvae develop on human corpses in predictable manner they are part of a forensic entomologist's toolbox for defining some portion of the post mortem interval.

Scientific investigation of *Chrysomya* would be aided by several kinds of genomic information. One would be the sequence of entire mitochondrial DNA (mtDNA) molecules. mtDNA is the locus of choice for identifying forensic insect evidence (Wells and Stevens, 2008), but this usually involves only short regions that may fail to resolve close relatives (*e.g.*, Wells et al. 2007). Short mtDNA regions have also been used for pure phylogenetic analysis, but resolving deeper divergences has also required labor intensive development of nuclear loci (*e.g.*, Singh and Wells 2011). It may be, however, that phylogenetic analysis of the

complete protein coding mtDNA sequence can work as well as one using nuclear DNA (Cameron et al. 2007). Another need is for more and better population genetic tools for forensically important flies such as *C. megacephala* (Wells and Steven 2008). Although some work has focused on amplified fragment length polymorphism (Bao and Wells 2014), it would be useful to have Mendelian genotypes such as microsatellites (Vasquez *et al.*, 2000; Ellegren, 2004; Farncombe *et al.*, 2014). The purpose of your study was to use a relatively small amount of NGS to develop new population genetic tools and data to assemble the mitochondrial DNA molecule.

2. SEQUENCING

2.1 ION TORRENT PGM

Genomic Extraction: Specimens

Chrysomya pacifica was caught by Dr. Jeffrey Wells on the island of New Britain in 1995. The whole fly was stored in 95% ethanol for around 6 months before extraction (Jeffrey Wells, personal communication). *Chrysomya megacephala* was caught in Southern Florida, USA by Josh Smith in 2013. They were also stored in 95% ethanol before extraction (Josh Smith, personal communication). All specimens were identified using morphology (Wells & Kurahashi, 1996).

Genomic Extraction: Methods

Chrysomya pacifica DNA was extracted from thoracic tissue using the phenol chloroform method. The whole genomic DNA was stored in -80°C freezer until 2014 (Jeffrey Wells, personal communication). The *C. megacephala* DNA samples were extracted using Qiagen's DNEasy Kit with samples derived from the abdominal tissue (Josh Smith, personal communication).

Results

The loading, quality, and polyclonal issues represented in Table 2.1.1 affected the yield and Ion Torrent PGM produced roughly 4.7 million reads. In theory, the 318 chip can produce 1.2 – 2 gigabases (Gb) (LifeTechnologies.com). Only 40.0-43.5% of the wells yielded usable reads. During subsequent data analysis, I applied a trimming algorithm to the raw reads on the basis of a PHRED score of 20 and ambiguity. For the ambiguity trim, I had the software remove all sequences that contained more than two consecutive ambiguous bases or Ns. This step further trimmed around 95% of the reads, thereby removing some of the reads and shortening the vast majority that remained. The analysis will be discussed later in this section.

$$2) \frac{\text{Total Number of Reads used in Assembly} * \text{Average Length of Reads}}{\text{Expected Genome Size of the Organism}}$$

Although no nuclear reference genome exists for neither *Chrysomya megacephala* nor *C. pacifica*, the size is likely around 650 million bases like other *Chrysomya* species (Picard *et al.*, 2012). If this is the case, the average coverage

TABLE 2.1.1: OUTPUT INFORMATION ON SUCCESSFUL CMOS WELL USAGE

Specimen	Input (ng/μL)	Loaded Wells (%)	Poly- clonal Wells	Quality	Usable ISPs (%)	Total Wells (%)	Reads Generated	Mean Read Length	Median Read Length	Mode Read Length
38 <i>Chrysomya pacifica</i>	9.4	70.3	21.0	16.9	62.1	43.5	4,916,088	278bp	306bp	379bp
39 <i>Chrysomya pacifica</i>	9.8	70.8	21.0	19.1	59.9	41.7	4,717,429	269bp	281bp	393bp
40 <i>Chrysomya megacephala</i>	9.9	70.3	23.3	19.3	57.3	40	4,542,261	291bp	338bp	383bp
41 <i>Chrysomya megacephala</i>	6.1	68.0	17.4	20.6	62.0	42	4,749,663	298bp	342bp	380bp

Input DNA quantification from Qubit in FIU DNA Core.

achieved by the Ion Torrent sequencing was around three fold (1.3 billion bases on average for each sample). Equation (2) above shows how to estimate the depth of coverage.

Most published reference genomes contain at least thirty-fold coverage (Li *et al.*, 2010; Sims, 2014). For the *Chrysomya* genus, one hundred fold coverage would require 650Gb of outputted sequencing data. The determination of a three-fold coverage suggests that the whole genome was characterized. In fact, the mitochondrial genome was overrepresented at more than one thousand fold coverage.

Ion Torrent is known for its inability to sequence homopolymer tracts (i.e. AAAAAAAAA). Sequence error rates rise dramatically after the seventh identical base incorporation (Quail *et al.*, 2012). After seven, the system failed to correctly include the proper amount of identical bases. It is as if it loses count of the homopolymers.

2.2 ILLUMINA MiSEQ

Genomic Extraction: Specimens

Two *Chrysomya megacephala* specimens were dissected and tissue from the head and abdomen was removed. On August 4th, 2014 Fanchen Bao and I dissected the specimen. The flies were caught in Southern Florida by Bao. They were classified by species and sex by morphology (Wells and Kurahashi, 1996).

Genomic Extraction: Method

The *C. megacephala* were extracted using Qiagen's DNEasy Kit with samples derived from the abdominal tissue

Results

The two genomic extractions were pooled in two MiSeq lanes on a flow cell and each yielded over 17 million raw sequencing reads.

2.3 CLC GENOMICS WORKBENCH ASSEMBLY

There are several values used to determine the quality of the assembly. First, the report shows how many contigs were generated. In theory, the number of contigs should be the number of chromosomes. In reality, the goal is to have a low number of contigs. There are several reasons why more contigs than chromosomes will be produced. Contamination from bacteria during library preparation can digest portions of the genome (Alkan *et al.*, 2010). The same contamination may produce extra reads from the bacterial genome. This is the case for several of the contigs produced in all the assemblies: bacterial genomic contributions from the *Wolbachia* genus were identified during BLAST searches of the contigs. Areas that are AT rich are notoriously difficult for all sequencers, especially Ion Torrent, and may not get sequenced (Quail *et al.*, 2012). Accurate sequencing reads with lower PHRED scores can be removed during the several trimming steps. All of these issues can hinder the assembly of whole chromosomes.

The most commonly used metric for assembly quality is the N50 value. CLC describes N50 as “the smallest of the largest contigs covering 50% of the total size of the contigs” (clcbio.com). The value provides more information than an average contig size and is a reliable piece of data used in bioinformatics (Cattonaro *et al.*, 2010; El-Metwally *et al.*, 2013). The N50 values for assemblies created with the Ion Torrent PGM sequencing data ranged from 568-607 and the values with Illumina MiSeq data were 1137 and 1143. The values indicate that the *de novo* assemblies created with Illumina data were of higher quality than those made with Ion Torrent data. The Illumina data also show relatively fewer contigs when the number of reads used is considered. The Illumina output was about 15 million reads which were used to create about 445,000 contigs. The Ion Torrent output was less than five million reads and yet around 245,000 contigs were created. However, both assemblies produced a large amount of contigs, indicating any or all of the aforementioned issues could be at fault. If the number of contigs had been much smaller, it would be possible to manually edit the contigs into a more concise assembly. Table 2.3.1 shows the metrics for all six assemblies.

TABLE 2.3.1: ASSEMBLY SUMMARY REPORT METRICS

		Total reads	Reads match	Contigs	N50	Min Length	Max Length	Average Length	Total bases
Ion Torren t PGM	38	4910753	3798555	241624	607	17	14935 ¹	565	136564028
	39	3905635	2982593	216207	568	42	14928 ¹	534	115442799
	40	4565304	3417447	246071	607	17	14931 ¹	565	139007495
	41	4745567	3702114	279274	620	15	14931 ¹	574	158467461
Illumin a MiSeq	42	16991702	15753778	443717	1143	200	112165	790	350575382
	43	16330718	14940015	446563	1137	200	112060	784	350199412

¹: The longest contigs from the Ion Torrent PGM assemblies represent the vast majority of the mitochondrial genome.

3. MICROSATELLITES

3.1 MSATCOMMANDER

I utilized two approaches: MsatCommander software, where I specified only the type of repeat such as dinucleotide, trinucleotide, and tetranucleotide ([https://code.google.com/p /msatcommander/](https://code.google.com/p/msatcommander/)), and CLC Bio Genomics Workbench, where I searched each contig for specific repeat motifs such as AAC or ATGG. The MsatCommander is an open source software program made available through research by Faircloth (2008). It searches inputted FASTA files for the user selected repeat options. See Figure 3.1.1. It uses Primer 3 software to develop the primers offered in the Primer Design option (Rozen & Skaletsky, 1999).

FIGURE 3.1.1: MSATCOMMANDER INTERFACE

(Image: Lauren Evenstone)

The software creates separated text files showing the sequence within the given contig. The sequence is annotated with the best options for forward and reverse primers, the repeat motif the primers define, and the amplification parameters necessary for a successful PCR. Each file also contains lesser options for the primer set. By analyzing the primers and the proposed amplicon, the software gives the suggested annealing temperature as well as any issues with self-complementarity that would impede proper amplification. Often, the repeat motif would not be an exact motif. As an example, CAACAACAACAACAA would be CAA repeated 5 times. Some of the results from using MSAT Commander had deviations such as is found in Figure 3.1.2.

FIGURE 3.1.2: EXAMPLE OF MSATCOMMANDER PRIMER 3 OUTPUT.

```
Using mispriming library misprime_lib_weight
Using 0-based sequence positions
OLIGO      start   len     tm       gc%    any     3'     rep seq
LEFT PRIMER      1899    23     60.14   47.83  4.00   2.00   6.70
TGATAAGACCTCTAGTGCAGAGC
RIGHT PRIMER      2142    21     58.29   47.62  5.00   2.00   5.36
TG TACTGGCAGGAATATGAGC
SEQUENCE SIZE: 2200
INCLUDED REGION SIZE: 2200

PRODUCT SIZE: 244, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 2.00
TARGETS (start, len)*: 2080,35


1860 AATGGATCTAAATATGTTTGACC AACATT TAAGTGAATTTGATAAGACCTCTAGTGCAGA
                                     >>>>>>>>>>>>>>>

1920 GCTTAACAATTATGGAGCAGCTGAAAAACACAACACTTGACTTGCAATTTGAAAACAAG
>>

1980 ACCTATAGCTTCTATGGCCTCACATGTTTTAATAAATGAAGATGACAAAGACAAGCAAAG

2040 GCAAAAAGCATACAAATCAGAACAGCTAAACACCTACAGCACCACCACCTCCTCCTCC
                                   *****

2100 TCCTCGGCCACCTCCACCACCTGCTCATATTCCTGCCAGTACAATGCTTAAACAAAGCAA
*****<<<<<<<<<<<<<<<<<<<

2160 ACAAGAAAAACAAAAGACCATCATCATCCTCAACATGTA
```

In this example, a CCT trinucleotide repeat, labeled with asterisks below the sequence, contains more than just CCT repeats. The sequence with forward facing arrows beneath it is the forward or left primer and the other with backward arrows is the reverse or right primer. The program provides location information, melting temperature (T_m), GC content, sequence length, and scores for self-complementarity for the primers.

3.2 CLC GENOMICS WORKSHOP

Using published information on insect microsatellites, I searched for the various motifs that have been successful with *C. megacephala* as well as other members of the family Calliphoridae (Florin & Gyllenstrand, 2002; He *et al.*, 2007; Torres *et al.*, 2008; Chong *et al.*, 2014; Farncombe *et al.*, 2014). The CLC also uses Primer 3 parameters to determine possible primer pairs, their annealing temperatures, GC content, and self-complementarity. Another similarity between the MsatCommander and the CLC is that both determine the primer pairs on the basis of individual contig information. The primers, however, are intended to be used on more than one individual and should only prime one location in the genome. Using CLC, I aggregated all contigs for all four *C. megacephala* and both *C. pacifica* for each repeat motif. I aligned the repeat sequences and designed primers in the flanking regions that would only amplify once per individual. That is to say that the primer pair sequences only appeared in one contig per specimen, or within a realistic amplifiable distance. Any contigs that contained both primers in the correct order were considered. The proposed

amplicon should be longer than 50bp (to prevent confusion with possible primer dimers) and shorter than 500bp. The maximum length was inferred from the short cycle length of the thermocycler. Cycle information and amplicon lengths will be discussed in PCR Protocol Development. Amplification results are shown in Appendix A and discussed later. Any product outside of the size range was disregarded.

The primer design portion of my project used only the Ion Torrent sequencing data as it was done before the Illumina MiSeq data were obtained. Later analysis on the Illumina data showed that the repeat motifs initially chosen were highly representative.

Aside from traditional STR loci, I also created Inter Simple Sequence Repeat (ISSR) markers. Research by Chong *et al.* (2014) suggested that since the *Chrysomya* genomes tend to be highly repetitive, using the repeat motif as the priming region could create reliable polymorphic loci. These markers would again only be counted if the resulting amplicon was longer than 50 and shorter than 500 base pairs long. The minimum length was especially important in the ISSR primer sets since the chance of primer dimer formation is inherently higher. The primer sets I designed can be found in Table 3.2.1.

TABLE 3.2.1: PRIMER SETS WITH OBSERVED AND EXPECTED SIZE RANGES

Primer Name ^α	Primer Sequence (5'-3')	T _m (°C)	Representative amplicon(bp) ^β	Observed range (bp) ^γ	Expected range (bp) ^δ	Motif
40-10017S	F:GAAGAGCAGAACGAGAAA	49.3	93	~150	290	(AAAT) _n
	R:ACCTACTACTATCTCTACC	46.0				
40-39693I	F:CCAACACAAACAACAACAA	50.8	234	~75, ~250	-	AAC
	C R:CTGTTGTGGTTGTTGTTGT	51.2				
40-15840S	F:CACCCCATCAATTACAAC	48.2	97	~75-150	82-97	(AAC) _n
	C R:CGACTAAAACCTACTTCAA	48.3				
40-39693S	F:AGAAATAAACAGGGCGCA	52.0	290	~315	213-239	(AAC) _n
	R:TGATGGTGGTGTGTTTTG	50.5				
38-7057I	F:GTTGTTGTTGTTGTTGTCC	50.0	423	~75, ~150	-	AAC
	R:GAACAACAACAACAACCG	49.7				
38-28185I	F:CAAGAACAACAACAACC	50.3	115	~75,~150, ~300	-	AAC
	C R:GTTGTTGTTGTTGTTGTTG	51.3				

α: S indicates an STR primer set where the primers flank a microsatellite locus. I indicates an ISSR primer where the primers are the repeat motifs. β: Representative repeat amplicon determined from Primer 3 analysis in CLC Bio. The specific contig that contained this amplicon is coded in the primer name (specimen 40, contig 10017, e.g.). γ: Observed following gel electrophoresis on 1% agarose (SeaChem) or 3% MetaPhor. δ: Lengths represented here are from searching all sequencing data for contigs that contain both forward and reverse primers. ISSR Primer sets were not included in expected range size as the unknown variable is the length of sequence between two microsatellites. F: Forward. R: Reverse.

3.3 PCR DEVELOPMENT

Methods

I began with the concentrations suggested by the Promega PCR Master Mix protocol which are shown in Table 3.3.1.

TABLE 3.3.1: SUGGESTED CONCENTRATION AND VOLUMES FOR PCR MASTER MIX

	Concentration	Hypothetical Volume (25 μ L total)	Actual Volume (25 μ L total)
PCR Master Mix	1x	12.5	12.5
Primer 1	0.25-1.0 μ M	0.5-2.5	2.5
Primer 2	0.25-1.0 μ M	0.5-2.5	2.5
PCR Grade Water	Variable	Variable	5.5
Input DNA	<250 ng	1-5	2

(Source: www.promega.com)

The wild caught samples came from extractions performed by Fanchen Bao using the Qiagen DNEasy Blood and Tissue Kit. In sets of five samples, I prepared and amplified 76 samples using six different primer pairs in singleplex reactions, 486 reactions in total including the negative controls. Thirteen of these sets were run electrophoretically on 1% SeaChem agarose gels. I manipulated the concentrations of the components. The ingredients and their volumes for each attempted reaction mix can be found in Table 3.3.2. . The final set included eight samples and two blanks using only two of the primer pairs and was run on an ABI Prism 3130 Capillary electrophoresis (Applied BioSystems).

TABLE 3.3.2: PCR MASTER MIX RECIPES

	S1 (μL)	S4 (μL)	S5 (μL)	S12 (μL)
Master Mix	12.5	12.5	14	13
Primer 1	2.5	2.5	3.0	3
Primer 2	2.5	2.5	3.0	3
PCR Grade Water	5.5	4	2	2.5
Input DNA	2	3.5	3	3.5

Amplification did occur for every reaction mix. However, not all primer pairs worked for each recipe. The thermocycler's stages, temperatures, and length of stages of the cycles can be found in Table 5.3.3 below (Labnet MultiGene Gradient PCR Thermal Cycler).

TABLE 3.3.3: THERMOCYCLER SPECIFICATIONS

	25 cycles					
	Initial	Denature	Anneal	Extend	Final	Hold
Temperature (C)	94	94	46-52*	72	72	4
Time	5 min	30 sec	30 sec	30 sec	10 min	∞

*The annealing temperature differed between the 12 rows in a gradient from 46°C to 52°C.

Since the primer pairs had a range of annealing temperatures, the gradient feature was used on the thermocycler. The twelve different rows were programmed to be different temperatures during the annealing stage of the cycles. The reaction tubes were positioned beforehand in the rows where their lowest annealing temperature would occur. The temperature in the gradient and the positioning of the primers can be found in Table 3.3.4. The thermocycler has twelve rows of eight wells. Primer sets that were placed into rows with slighter

higher than optimal annealing temperatures amplified just as consistently as those positioned in lower than optimal or optimal temperatures.

TABLE 3.3.4: POSITIONING OF THE SINGLEPLEX REACTIONS WITHIN THE THERMOCYCLER

Primer Pair	Row	Temp. (c)	Lower Tm*
40-10017S	1	46.0	46 (Reverse)
40-15480S	5	48.3	48.2 (Forward)
38-7057I	7	49.6	49.7 (Reverse)
38-28185I	7-8	49.6-50.3	50.3 (Forward)
40-39693S	8-9	50.3-51.0	50.5 (Reverse)
40-39693I	9	51.0	50.8 (Forward)

*Annealing temperature (Tm) of the primer within the primer set that is the lower of the two. If the reverse has 46C and the forward has an annealing temperature of 49.3C, the tubes were positioned to experience the lower of the two, or 46C.

The resulting bands developed by gel electrophoresis are presented in Table 3.3.5. Photographs of the gels themselves can be found in Appendix A.

These gels follow the optimization of the PCR reaction mix and gel conditions. All gels contain a 1kb Ladder (N3234L, New England Biolabs, Ipswich, MA). Final sets were conducted on 3% MetaPhor gels. They contained the same ladder as well as a 20bp Molecular Ruler (#170-8201, Bio-Rad Laboratories, Hercules, CA).

Results

No clear evidence of polymorphism was apparent on the gels. The 1% SeaChem agarose gels produced blurry and faint bands. A 3% MetaPhor agarose gel and a longer electrophoresis system were employed to increase resolution. These high density gels failed to produce clear results. Two primer

pair sets, 40-15840S and 38-28185I, which showed possible potential over several reaction mixes, were chosen to be reordered as fluorescently labeled primers for use in capillary electrophoresis. The primer sets were modified on the 5' end with 6-FAM via N-hydroxysuccinimide esterification on the oligonucleotide's phosphate group. Using an ABI Prism 3130 Genetic Analyzer, the amplified samples were combined with a LIZ labeled internal size standard and highly deionized formamide at the Forensic DNA Profiling Facility at Florida International University.

The preliminary results for the STR primer show homozygous and heterozygous peaks across a small range of sizes, from 82-97 bases. These sizes correlate with the sequence data and the range of sizes seen in the different assemblies (see Table 3.2.1). In each electropherogram, there is pronounced stutter or artifacts. These were determined not to be allele peaks because their heights were less than 10% of the dominant peak or peaks.

There is no relationship information available on the samples used for this project. Further studies into the STR primer set should include known parents and offspring to determine if this locus is inherited via Mendelian genetics. The electropherograms for this primer set can be found in Appendix E. Preliminary results for the ISSR primer set contained many more peaks that were expected from the agarose gels. They will require more work in future studies to make them legible.

TABLE 3.3.5: RESULTS OF SINGLEPLEX PCR BY AGAROSE GEL ELECTROPHORESIS

Reaction mix		S1			S4	S5			
Tm	Primers	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8
		78-82	83-87	88-92	93-97	98-102	103-107	108-112	113-117
47.6 5	40- 10017S	~300F	None	None	75-300S	None	~250F	None	None
51.0	40- 39693I	None	250V	200V	50-300S	PD	375	None	250V, 400V
48.2 5	40- 15840S	None	83 & 85: 100-200S	~100: 88, 89, 91, 92	75-150FS	None	None	None	None
52.0	40- 39693S	None	83 & 85 250-300S 84, 86, 87: 350F	300S: 88, 89, 91, 92	300S	~300F	None	~800- 1000FV	100S, 250
49.8	38- 7057I	~100V	83:250S 84, 85,	~50- 250S: 89, 91, 92	75-200VS	None	None	~1000F V	~300
50.8	38- 28185I	75- 350S	50-300S	45-250V: 88, 89, 91, 92. C!	50-300VS	~150F	None	~1000F	200- 300VS

F: Faint. V: Variable. S: Smear. PD: Only presumed primer dimers present. C!: possible contamination indicated by banding in PCR blank.

Table Continues on next page*

TABLE 3.3.5: RESULTS OF SINGLEPLEX PCR BY AGAROSE GEL ELECTROPHORESIS (CONTINUED)

Reaction mix		S5		S4	S12	
Tm	Primers	Set 9	Set 10	Set 11	Set 12	Set 13
		118-122	123-127	128-132	133-137	138-142
47.65	40-10017S	v	~150F	None	None	-
51.0	40-39693I	PD	50PD, ~200	50-300FS	None	-
48.25	40-15840S	None	None	None	None	None
52.0	40-39693S	275	~250: 123, 126, 127	None	None	None
49.8	38-7057I	None	None	None	~150	-
50.8	38-28185I	75F, 250F, 1000F	None	None	50-300S	-

F: Faint. V: Variable. S: Smear. PD: Only presumed primer dimers present. C!: possible contamination indicated by banding in PCR blank

4. MITOCHONDRIAL GENOME

4.1 ASSEMBLY

My results were assembled by mapping *Chrysomya pacifica* contigs to published *C. megacephala* mitochondrial genomes JX913738 and JX913739 (Nelson et al. 2012). Both were made available from the research performed by Nelson et al. (2012). Using JX913738, which I will refer to as Mega1 in the results, I created an alignment in CLC with each of the suspected contigs. I did the same with JX913739, which I will refer to as Mega2. Although a few conflicting bases were evident, they contained the coding region and short intergenic non-coding regions of the mitochondrial genome. They did not have the variable region. Also called the D-Loop or the Control Region, there is a string of bases that vary between members of the same species. The rest of the genome is highly conserved within species (Russell, 2006).

Using CLC, I used the individual assemblies to create internal BLAST databases in order to “Reverse-BLAST” the published genome against the contigs. The program BLAST normally works by having a database of known sequences that you search with your unknown sequence. The Reverse-BLAST uses the known sequence to find similar sequences within the hundreds of thousands of sequences. I used adjusted parameters to allow for less than perfect identity which was the initial issue with simply searching the sequences. The results of each Reverse-BLAST showed the sequences and how they aligned with Mega1 and later Mega2. It also showed where point differences

occurred. I manually assembled the few contigs into the mitochondrial genome using the coding information. *Chrysomya pacifica* does not have any published genome information. The only published sequence data is that of the cytochrome oxidase subunit 1 used in mitochondrial DNA barcoding, as seen on BOLDSystems.org (Ratnasingham & Hebert, 2007). I present in Appendix B the full annotated mitochondrial genome for *C. pacifica*. In Appendix C, a comparison of *C. pacifica*, Mega1, and Mega2 with translational divergences can be found. Table 4.1.1 below shows the order of the coding sequences found in the *C. pacifica* mitochondrial genome.

TABLE 4.1.1: CODING SEQUENCES WITHIN THE *CHRYSOMYA PACIFICA* MITOCHONDRIAL GENOME

Start	Strand	Type	Name	Anticodon
1..66 ³	+ ⁴	tRNA	Ile Isoleucine	31..33 (gat) ⁵
136..71	-	tRNA	Gln Glutamine	109..107 (ttg)
148..213	+	tRNA	Met Methionine	178..180 (cat)
217..>1231	+	CDS	ND2 NADH Dehydrogenase subunit 2 ¹	
1232..1299	+	tRNA	Trp Tryptophan	1262..1264 (tca)
1355..1292	-	tRNA	Cys Cysteine	1324..1326 (gca)
1429..1363	+	tRNA	Tyr Tyrosine	1396..1398 (gta)
<1428..>2961	+	CDS	COX1 Cytochrome Oxidase Subunit 1 ¹²	
2962..3027	+	tRNA	Leu Leucine	2991..2993 (taa)
3033..>3420	+	CDS	COX2 Cytochrome Oxidase Subunit 2 ¹	
3721..3791	+	tRNA	Lys Lysine	3751..3753 (ctt)
3791..3857	+	tRNA	Asp Aspartate	3822..3824 (gtc)
3858..4022	+	CDS	ATP8 ATP Synthase F0 Subunit 8	
4016..4693	+	CDS	ATP6 ATP Synthase F0 Subunit 6	
4698..5486	+	CDS	COX3 Cytochrome Oxidase Subunit 3	
5493..5557	+	tRNA	Gly Glycine	5523..5525 (tcc)
5558..5911	+	CDS	ND3 NADH Dehydrogenase subunit 3	
5914..5978	+	tRNA	Ala Alanine	5943..5945 (tgc)
5978..6041	+	tRNA	Arg Arginine	6007..6009 (tcg)
6048..6113	+	tRNA	Asn Asparagine	6078..6080 (gtt)
6114..6181	+	tRNA	Ser Serine	6139..6141 (gct)
6183..6249	+	tRNA	Glu Glutamate	6213..6215 (ttc)
6334..6268	-	tRNA	Phe Phenylalanine	6302..6300 (gaa)

8054..>6334	-	CDS	ND5	NADH Dehydrogenase Subunit 5 ¹	
8134..8070	-	tRNA	His	Histidine	8104..8012 (gtg)
9473..>8135	-	CDS	ND4	NADH Dehydrogenase Subunit 4 ¹	
9763..9467	-	CDS	ND4L	NADH Dehydrogenase Subunit 4L	
9766..9830	+	tRNA	Thr	Threonine	9796..9798 (tgt)
9896..9831	-	tRNA	Pro	Proline	9866..9864 (tgg)
9899..10423	+	CDS	ND6	NADH Dehydrogenase Subunit 6	
10423..>11557	+	CDS	CYTB	Cytochrome Oxidase B*	
11558..11325	+	tRNA	Ser	Serine	11588..11590 (tga)
12580..11642	-	CDS	ND1	NADH Dehydrogenase Subunit 1	
12655..12591	-	tRNA	Leu	Leucine	12626..12624 (tag)
13984..12656	-	rRNA	16s	16s large ribosomal RNA	
14056..13985	-	tRNA	Val	Valine	14023..14021 (tac)
14840..14057	-	rRNA	12s	12s small ribosomal RNA	
14841..15716	+	Misc.		Variable Region	

¹: TAA stop codon is completed by 3' adenylation of mRNA. ²: Nontraditional Start Codon. ³: Position number reflects Isoleucine as the first coding sequence. However, these numbers are technically arbitrary given that mitochondrial DNA is circular. ⁴: A positive sign (+) indicates that the coding sequence is on the strand complementary to sequence on the negative (-) strand. These indications are again, arbitrary but consistent. ⁵: The anticodon sequence located at the indicated position is presented within parentheses.

To confirm that this assembly was accurate, I used several open source web servers as well as CLC Genomics Workbench analyses for proteins, RNA, and nucleotides.

4.2 CLC GENOMICS WORKBENCH ANALYSIS

Once aligned to Mega1 and Mega2, it was easy to see where bases differed between the species and where insertions or deletions (indels) occurred. The published genomes contained annotations for coding regions, tRNAs and rRNAs. These coding regions of the two published mitochondrial genomes were then copied to the *C. pacifica* genome. To ensure that the genes and RNAs translated and folded in the same way in both species, I applied different tools to each part of the sequences. An alignment of both reference genomes and the *C. pacifica* genome can be found in Appendix C.

For the genes or CDS (Coding DNA Sequence), the sequence was translated in every possible reading frame using the amino acid code for invertebrate mitochondria. The published annotation for Mega1 and Mega2 were used as references. The *C. pacifica* and both of the *C. megacephala* mitochondrial genomes translated in the same reading frame for every CDS, meaning that for all six reading frames (+1, +2, +3, -1, -2, -3), only one reading frame did not contain stop codons in the middle of the sequence. In some instances no stop codon was detected. This stop codon can be resolved during transcription when the addition of a poly-A tail can occur (Nelson et al., 2012; Slomovic, et al., 2006) which can turn a terminal cytosine into the stop codon CAA. In some cases, a methionine does not start the translation. In these cases, an isoleucine or a serine acts as the start codon. The alternative start amino acids are also seen in the published mitochondrial genomes (Nelson et al., 2012; Slomovic, et al., 2006).

In five of the CDS regions, the differences in nucleotide sequence caused a change in the amino acid. The majority of the conflicts between *C. pacifica* and the published *C. megacephala* data were silent changes that did not result in amino acid changes. For these five genes, ND1, ND5, ND4, CYTB, and COX1, protein analysis tools provided in the CLC software were employed. A table of all base differences between *C. pacifica* and the two *C. megacephala* references can be found in Table 4.2.1. Very few base differences result in a changed amino acid, indicated in the column labeled “AA Change.” Since Mega1 and Mega2

differ from one another, both are considered here. Also, the eight changes or replacements that occur do not distinguish *C. pacifica* from the *C. megacephala*. Of the 108 differences listed in Table 4.2.1, only 18 are unique to *C. pacifica* and all of them are silent. The region labeled trnL for leucine that does not occur within the *C. pacifica* genome will be discussed later. Those base differences within the region designated “VR” are within the non-coding variable region that occurs between the 12s rRNA and the Leucine tRNA at position 1 when the genome is in its natural, circular formation.

TABLE 4.2.1: COMPARISON OF BASE DIFFERENCES BETWEEN *C. PACIFICA* AND TWO *C. MEGACEPHALA* REFERENCE GENOMES.

Pos- ition	Mega1	Pacifica	Mega2	CDS	AA
396	T	T	C	ND2	
519	G	A	G	ND2	
549	T	T	C	ND2	
891	C	C	T	ND2	
1356	-	-	C	INT	
1360	G	G	A	INT	
1474	T	T	C	COX1	
1785	G	A	G	COX1	
2070	T	T	C	COX1	
2085	T	C	T	COX1	
2115	T	C	C	COX1	
2214	A	G	G	COX1	
2241	G	G	A	COX1	
2641	A	A	G	COX1	S S G
2646	G	G	A	COX1	
3235	C	C	T	COX2	
3429	C	C	T	COX2	
3892	T	T	C	ATP8	
3984	T	T	C	ATP8	
4034	T	T	C	ATP6	
4205	T	T	C	ATP6	
4286	C	T	C	ATP6	
4920	C	T	C	COX3	
4953	C	C	T	COX3	
5040	A	A	G	COX3	
5298	G	A	A	COX3	
5783	T	A	A	ND3	
6046	T	-	-	INT	
6239	T	-	-	trnE	
6357	G	A	A	ND5	
6498	A	A	G	ND5	
6665	C	C	A	ND5	V V L
6666	C	C	T	ND5	V V L
6844	A	G	G	ND5	F S S
6855	G	A	A	ND5	
6918	G	G	A	ND5	
7125	C	C	T	ND5	
7233	A	A	T	ND5	
7476	T	T	C	ND5	
7743	T	T	C	ND5	
7770	C	C	T	ND5	
7845	G	G	A	ND5	
8018	A	G	G	ND5	
8188	A	G	G	ND4	
8247	C	T	T	ND4	
8316	T	C	C	ND4	
8385	T	T	C	ND4	
8388	C	C	T	ND4	
8472	T	T	C	ND4	
8535	T	T	C	ND4	
8592	T	T	C	ND4	
8706	T	T	C	ND4	
8745	C	C	T	ND4	
8775	C	T	T	ND4	
8783	C	C	T	ND4	
8786	C	C	T	ND4	

8877	A	G	A	ND4	
8949	G	A	A	ND4	
9006	T	T	A	ND4	
9174	C	C	T	ND4	
9312	A	A	G	ND4	
9342	A	G	G	ND4	
9384	T	T	C	ND4	
9425	A	G	G	ND4	Y H H
10033	A	A	G	ND6	
10180	T	C	T	ND6	
10274	T	T	C	ND6	
10638	C	T	C	CYTB	
10754	A	T	T	CYTB	Y F F
10821	C	C	T	CYTB	
10905	C	C	T	CYTB	
10917	A	G	A	CYTB	
10956	C	T	T	CYTB	
11097	T	T	C	CYTB	
11202	A	A	G	CYTB	
11211	T	A	T	CYTB	
11289	A	G	G	CYTB	
11314	C	C	T	CYTB	
11358	T	C	T	CYTB	
11431	C	T	T	CYTB	
11433	G	G	A	CYTB	
11497	G	G	A	CYTB	V V I
11514	C	C	T	CYTB	
11657	C	C	T	ND1	
11756	C	C	T	ND1	
11795	C	A	A	ND1	
11891	T	T	C	ND1	

11963	T	T	C	ND1	
12028	A	C	C	ND1	S A A
12320	T	T	C	ND1	
12760	G	-	G	16s	
12844	C	T	T	16s	
13273	C	A	C	16s	
13534	C	T	T	16s	
13737	C	T	T	16s	
14306	C	T	T	12s	
14307	A	C	C	12s	
14522	C	T	T	12s	
14849	A	C	C	INT	
14935- 15003		-		trnL2	NOT in pacific
15059	T	T	C	VR	
15094	C	C	T	VR	
15103	T	C	T	VR	
15128	A	G	A	VR	
15139	C	C	T	VR	
15167	C	T	C	VR	
15182	T	C	T	VR	
15194	C	C	T	VR	
15698	-	C	A	VR	
15749	-	A	G	VR	

INT: Intergenic.
VR: Variable Region

CLC Protein Analysis

For the five gene coding regions that are different from the reference genomes, I employed the protein analysis tools offered in CLC Genomics Workbench. The genes- ND1, ND4, ND5, CYTB, and COX1 were translated using CLC to simulate a protein product. Each protein was assessed on the basis of its translated sequence, hydrophobicity, charge, protein domains, and transmembrane predictions

Translation

Invertebrate mitochondrial RNA is translated using a slightly different amino acid code. In the invertebrate mitochondrial translation code there are two traditional stop codons: UAA and UAG. Five of the thirteen proteins coded in the genome require modification to the mRNA to achieve a stop codon. By the addition of a poly(A) tail, a terminal U or UA can become UAA and successfully stop translation (Slomovic *et al.*, 2006).

Sequences are started by both methionine and isoleucine. This indicates that only the first two letters of the codon are important because methionine is AUA and AUG and isoleucine is AUU and AUC. The cytochrome oxidase subunit 1 begins with a serine. A study on seven beetle species' COX1 gene shows that the translation product begins with several non-traditional start codons (Sheffield *et al.*, 2008). They found that within the Coleoptera order, leucine, arginine,

asparagine, glutamate, lysine, and isoleucine act as start codons. In the case of *C. pacifica*, the start codon is a serine.

During translation, all open reading frames were considered. The differences in sequence between *C. pacifica* and the *C. megacephala* references did not result in any frame shifts.

Hydrophobicity

I searched the five proteins in question for stretches of hydrophobic amino acid residues. The plots are shown below in Figure 6.2.1. Each plot is an overlay of all three mitochondrial genomes. The blue line is hydrophobicity of the amino acids as a function of position within the *C. pacifica* sequence. Any deviations from this line caused by the difference in amino acid sequence are visible as a green line (Mega1) or a red line (Mega2). The hydrophobicity charts showed minimal changes across all three genomes for the five proteins (Figure 6.2.1a-e). Hydrophobic stretches of the poly peptides form beta strands and are conducive to residing within a lipid bi-layer membrane. These proteins are a part of the respiration pathway within the mitochondrion. This pathway uses a proton gradient across the membrane of the cristae to produce cellular power in the form of ATP (Nicholls, 1974).

FIGURE 4.2.1: HYDROPHOBICITY OVERLAYS OF *C. PACIFICA*, MEGA1, AND MEGA2

a)

b)

c)

d)

e)

Hydrophobicity plots for a) ND1, b) ND4, c) ND5, d) CYTB, and e) COX1. Y axis is a hydrophobicity scale and X axis is the amino acid position within the protein. *Chrysomya pacifica* is blue, Mega1 is green, and Mega2 is red.

Charge

The change in the amino acid sequence did result in a minor change in the overall charge of the transcribed protein. The change in charge, however, only occurs outside physiological pH in the acidic and basic extremes (Llopis *et al.*, 1998)

Transmembrane Helix Prediction

To identify the transmembrane regions of proteins, CLC connects to the Center for Biological Sequence Analysis website to use the TMHMM Server version 2.0 (Moller *et al.*, 2002). This website locates helix and beta strand structures inferred by the amino acid sequence to predict where transmembrane proteins would pass through the hydrophobic inside of the lipid bi-layer.

Hydrophilic spans of the amino acid sequences are labeled as either inside or outside the membrane. This analysis was performed on ND1, ND4, ND5, CYTB, and COX1 to see if the translational differences between *C. pacifica* and the two *C. megacephala* cause a change in the transmembrane structure of the final, three dimensional protein. A change in this structure could mean a change in the mitochondrion's ability to perform respiration. It was determined that the change in amino acid did not prevent the three NADH dehydrogenase subunits, the cytochrome oxidase B, nor the cytochrome oxidase subunit 1 proteins from spanning the membrane in a similar way as found in the other organisms.

Pfam

Using the Protein Family database Pfam, (version 27), I performed an analysis on these five proteins to assess if changes in the amino acid structure affected any protein domains. Hidden Markov Model profiles for available protein sequences were compared to my submitted proteins. All five sequences returned

identical Pfam domains as the *C. megacephala* reference. See Table 6.2.2 for domains. The Mega2 reference genome only differed once with the *C. pacifica* genome within these five proteins. Therefore, it is only shown on the Pfam table for the cytochrome oxidase subunit 1 protein domains. The amino acid differences between the references and the novel genome did not occur in any of these domains. It is consequently assumed that the amino acid differences do not affect the activity of the protein.

tRNA and rRNA Folding Parameters

There are several different functions that RNA provides within the cell. The ribosomal RNA (rRNA) is the structure within which messenger RNA (mRNA) is translated and the transfer RNA (tRNA) provides amino acids. This process produces a polypeptide chain, the most basic form of a protein. Found in Appendix D, I have assessed each tRNA sequence for a balance between the lowest free energy structure and a useful structure. The decision to proceed with a higher free energy form determined from the reference genomes and visual inspection using CLC Genomics Workbench. Only one tRNA, glutamate, differed between *C. megacephala* reference genomes and the *C. pacifica* genome. The reference genome designated Mega1 has a thymine at position 6239 where Mega2 and *C. pacifica* do not. The insertion/deletion (indel) does not affect the hairpin loop responsible for acting as the anticodon.

The ribosomal RNA (rRNA) folding was much harder to determine. The long and short rRNAs, or 16s and 12s, respectively, are much more complex

than the tRNAs. The published reference genomes lack information on RNA structure. A comparison of the output from CLC Genomics Workbench and other open source programs can be found in the next section.

4.3 OPEN SOURCE ANALYSIS

The assembled mitochondrial genome for *C. pacifica* was submitted to the MITOS Webserver (<http://mitos.bioinf.uni-leipzig.de/index.py>). MITOS uses BLASTX searches for varying lengths and positions within the submitted mitochondrial genome (Bernt *et al.*, 2013). MITOS generated a table containing each CDS, tRNA, and rRNA with their corresponding position within the mitochondrial genome, which strand they code from, and the secondary structures of the RNAs. For several of these features, the position (and therefore the sequence) differed from that of the published data. MITOS was consistent, however, with its annotation. I submitted both Mega1 and Mega2 and the results were consistent with the *C. pacifica* mitochondrial genome rather than the published annotations. Below are the lengths given for each annotation for the three submitted genomes. The three columns under the MITOS header contain the output data from their webserver. The two under the GenBank header are the reference genomes whose annotations reflect published data. The CLC header contains the *C. pacifica* genome that was annotated by aligning it to the GenBank genomes.

TABLE 4.2.2: PROTEIN DOMAINS FOR ND1, ND4, ND5, CYTB, COX1

Organism	CDS	Domain	Position	Accession	Score	E-value	Description
Pacifica	ND1	NADHdh	5..306	PF00146.16	421.5	1.60E-126	NADH dehydrogenase
Mega1	ND1	NADHdh	5..306	PF00146.16	420.6	3.10E-126	NADH dehydrogenase
Pacifica	ND4	Oxidored_q1	106..373	PF00361.15	241.3	1.10E-71	NADH- Ubiquinone /plastoquin one (complex 1), various chains
		Oxidored_q5_N	1..103	PF0159.12	76.4	1.40E-21	NADH- ubiquinone oxidoreductase chain 4, amino terminus
Mega1	ND4	Oxidored_q1	106..373	PF00361.15	241.2	1.10E-71	NADH- Ubiquinone /plastoquin one (complex 1), various chains
		Oxidored_q5_N	1..103	PF0159.12	75.7	2.20E-21	NADH- ubiquinone oxidoreductase chain 4, amino terminus
Pacifica	ND5	Oxidored_q1	102..366	PF00361.15	257.4	1.30E-76	NADH- Ubiquinone /plastoquin one (complex 1), various chains
		NADH5_C	389..570	PF06455.6	217.7	7.00E-65	NADH dehydrogenase subunit 5 C-terminus
		Oxidored_q1_N	29..91	PF00662.15	55.9	2.20E-15	NADH- ubiquinone oxidoreductase chain 5, amino terminus
Mega1	ND5	Oxidored_q1	102..366	PF00361.15	257.1	1.50E-76	NADH- Ubiquinone /plastoquin one (complex 1), various chains
		NADH5_C	389..570	PF06455.6	216.4	1.80E-64	NADH dehydrogenase subunit 5 C-terminus
		Oxidored_q1_N	29..91	PF00662.15	55.9	2.20E-15	NADH- ubiquinone oxidoreductase chain 5, amino terminus

Mega2	ND5	Oxidored_q1	102..366	PF00361.15	257.4	1.30E-76	NADH-Ubiquinone /plastoquinone (complex I), various chains
		NADH5_C	389..570	PF06455.6	218	6.00E-65	NADH dehydrogenase subunit 5 C-terminus
		Oxidored_q1_N	29..91	PF00662.15	55.	2.20E-15	NADH-Ubiquinone oxidoreductase (complex I), chain 5 N-terminus
Pacifica	CYTB	Cytochrom_B_N_2	89..254	PF13631.1	200.3	1.80E-59	Cytochrome b(N-Terminal)/b6/petB
		Cytochrom_B_C	259..360	PF00032.12	140.7	1.50E-41	Cytochrome b(C-Terminal)/b6/petB
Mega1	CYTB	Cytochrom_B_N_2	89..254	PF13631.1	202.4	4.00E-60	Cytochrome b(N-Terminal)/b6/petB
		Cytochrom_B_C	259..360	PF00032.12	140.7	1.50E-41	Cytochrome b(C-Terminal)/b6/petB
Mega2	CYTB	Cytochrom_B_N_2	89..254	PF13631.1	200.3	1.80E-59	Cytochrome b(N-Terminal)/b6/petB
		Cytochrom_B_C	259..360	PF00032.12	140.7	1.50E-41	Cytochrome b(C-Terminal)/b6/petB
Pacifica	COX1	COX1	9..458	PF00115.15	565.8	5.30E-170	Cytochrom C and Quinol oxidase polypeptide I
Mega1	COX1	COX1	9..458	PF00115.15	565.8	5.30E-170	Cytochrom C and Quinol oxidase polypeptide I
Mega2	COX1	COX1	9..458	PF00115.15	564.4	1.40E-169	Cytochrom C and Quinol oxidase polypeptide I

TABLE 4.3.1: COMPARISON OF CODING REGION LENGTHS IN MITOS, PUBLISHED *C. MEGACEPHALA* GENOMES, AND *C. PACIFICA*

	MITOS			CLC	GenBank	
Name	Pac	Mega1	Mega2	Pac	Mega1	Mega2
trnI(gat)	66	66	66	66	66	66
trnQ(ttg)	69	69	69	69	69	69
trnM(cat)	69	69	69	69	69	69
nad2-0	882	882	882	1015	1015	1015
nad2-1	18	18	18	-		
trnW(tca)	68	68	68	68	68	64
trnC(gca)	64	64	64	64	64	64
trnY(gta)	67	67	67	67	67	67
cox1	1509	1509	1509	1534	1534	1534
trnL2(taa)	66	66	66	66	66	66
cox2	672	672	672	688	688	688
trnK(ctt)	70	70	70	70	70	70
trnD(gtc)	67	67	67	67	67	67
atp8	162	162	162	165	165	165
atp6	666	666	666	678	678	678
cox3	783	783	783	789	789	789
trnG(tcc)	65	65	65	65	65	65
nad3	348	348	348	355	355	355
trnA(tgc)	64	64	64	65	65	65
trnR(tcg)	64	64	64	64	64	64
trnN(gtt)	66	66	66	66	66	66
trnS1(gct)	68	68	68	70	70	70
trnE(ttc)	67	68	67	67	68	67
trnF(gaa)	67	67	67	67	67	67
nad5-0	1695	1695	1695	1720	1720	1720
trnH(gtg)	65	65	65	65	65	65
nad4	1332	1332	1332	1339	1339	1339
nad4l	261	261	261	297	297	297
trnT(tgt)	65	65	65	65	65	65
trnP(tgg)	66	66	66	66	66	66
nad6	498	498	498	525	525	525
cytb	1125	1125	1125	1134	1134	
trnS2(tga)	68	68	68	68	68	68
nad1	909	909	909	939	939	939
trnL1(tag)	65	65	65	65	65	65
16s	1363	1364	1364	1329	1330	1329
trnV(tac)	72	72	72	72	72	72
12s	786	786	786	784	784	784
nad5-2	333	66	66	-	-	-

TABLE 4.3.2: COMPARISON OF THE POSITIONS AND LENGTHS OF *C. PACIFICA* CODING REGIONS BETWEEN MITOS AND CLC

	MITOS			CLC		
Name	Start	Stop	Length	Start	Stop	Length
trnI(gat)	1	66	66	1	66	66
trnQ(ttg)	71	139	69	71	139	69
trnM(cat)	148	216	69	148	216	69
nad2-0	241	1122	882	217	1231	1015
nad2-1	1164	1181	18	-	-	-
trnW(tca)	1232	1299	68	1232	1299	68
trnC(gca)	1292	1355	64	1292	1355	64
trnY(gta)	1363	1429	67	1363	1429	67
cox1	1434	2942	1509	1428	2961	1534
trnL2(taa)	2962	3027	66	2962	3027	66
cox2	3033	3704	672	3033	3720	688
trnK(ctt)	3721	3790	70	3721	3791	71
trnD(gtc)	3791	3857	67	3791	3857	67
atp8	3858	4019	162	3858	4022	165
atp6	4019	4684	666	4019	4693	678
cox3	4698	5480	783	4698	5486	789
trnG(tcc)	5493	5557	65	5493	5557	65
nad3	5558	5905	348	5558	5911	355
trnA(tgc)	5914	5977	64	5914	5978	65
trnR(tcg)	5978	6041	64	5978	6041	64
trnN(gtt)	6048	6113	66	6048	6113	66
trnS1(gct)	6114	6181	68	6113	6182	68
trnE(ttc)	6183	6249	67	6183	6249	67
trnF(gaa)	6268	6334	67	6268	6334	67
nad5-0	6351	8045	1695	6334	8045	1720
trnH(gtg)	8070	8134	65	8070	8134	65
nad4	8142	9473	1332	8135	9473	1339
nad4l	9470	9730	261	9470	9730	261
trnT(tgt)	9766	9830	65	9766	9830	65
trnP(tgg)	9831	9896	66	9831	9896	66
nad6	9911	10408	498	9899	10423	525
cob	10423	11547	1125	10423	11557	1125
trnS2(tga)	11558	11625	68	11558	11625	68
nad1	11663	12571	909	11642	12580	939
trnL1(tag)	12591	12655	65	12591	12655	65
rrnL	12624	13986	1363	12656	13984	1329
trnV(tac)	13985	14056	72	13985	14056	72
rrnS	14055	14840	786	14057	14840	784
nad5-2	15362	15694	333	-	-	-

In the case of ND2 and ND5, the open source software determined the CDS were split. Table 4.3.1 shows the differences in position determined by MITOS and by the Nelson Group (JX913738 and JX913739). Table 4.3.2 shows the differences of both length and position for *C. pacifica*. The webserver produces as output a graph of coding region probabilities based on BLASTX searches. This MITOS graph, found below, shows the lower probabilities for the mistaken spliced CDS. The quality scores show eight of the thirteen genes as

FIGURE 4.3.1: QUALITY SCORES OF MITOS DESCRIBED CODING REGIONS

(<http://mitos.bioinf.uni-leipzig.de/index.py>)

being above 10^5 . One of the genes, ATP8, has a quality score that falls below three genes that were not called. MITOS does not offer information on threshold limits used in this graph. Both ND2 and ND5's suggested separate sequences are found below a quality score of 10. In addition to this graph, I used information

of the translated sequence to determine that the spliced proteins are not realistic. I rejected the positions of CDS offered in the MITOS output.

The mitochondrial genome is known for its density of coding sequence. In some cases, gene sequence boundaries are determined by the level of overlap with neighboring genes or RNAs (Sheffield *et al.*, 2008). If the MITOS system uses this same reasoning, the RNAs are likely inaccurate as well.

The difference in start and stop positions in the genes caused there to be different translations in the proteins and different folding in the RNAs. The largest difference between the CLC data and the MITOS data was the suggested folding of the large and small RNA. No reference data exists on how the publishers of Mega1 and Mega2 demonstrated RNA folding. Found in Appendix D, I present three different folded structures for 16s and 12s rRNA. In addition to the results produced by MITOS, I have included predicted secondary structures created by the LocARNA webserver (Smith *et al.*, 2010). LocARNA is made available by the Freiburg Bioinformatics Group; it uses Zuker's mfold algorithm to fold the sequences according to a realistic maximum free energy (MFE) conformation (Smith *et al.*, 2010). The article published in Molecular Phylogenetics and Evolution by Bernt and his peers in which MITOS is presented as a novel software option fails to explain its folding parameters (2013). Only CLC, which also uses Zuker's mfold algorithm, offered any information on what the folding energy was, which can be found in Appendix D (CLC Manual).

4.4 PHYLOGENETICS

Phylogenetic analyses were performed using sequencing data from both *C. megacephala* reference genomes, *C. pacifica* produced within this project and publically available sequencing data from other calliphorid species. Traditionally, the cytochrome oxidase subunits have been used to monitor evolution, because the rate of mutation in this section of the mitochondrial genome tends to be consistent (Brower, 1994). Bold Systems has created a database that houses information primarily on cytochrome oxidase subunit I for all available species; the region is considered the “Barcode of Life” (Ratnasingham & Hebert, 2007). In the last year alone, the site has had over 675,000 sequences uploaded (boldsystems.org). The region, however, may not offer universal genetic evolutionary information for the calliphorid family (Nelson *et al.*, 2012). Indeed, a pairwise comparison generated in CLC proves that cytochrome oxidase subunit 1 (1533b) differs from pairwise comparisons generated with whole mitochondrial genome data. These comparisons were made using the four *C. megacephala* specimens and two *C. pacifica* sequenced for this project. Additionally, I added publically available sequencing data from seven species of *Chrysomya* flies: *C. bezziana*, *C. albiceps*, *C. putoria*, *C. pinguis*, *C. saffrana*, *C. rufifacies*, and the *C. megacephala* mitochondrial genomes referred to as Mega1 and Mega2. Finally, as out-groups, I also included *Lucilia sericata* and *Cochliomyia hominivorax*. The publically available genomes are listed in Table 4.4.1.

TABLE 4.4.1: SOURCES OF SPECIES AND SPECIMENS USED IN PHYLOGENETIC ANALYSES

Species	Accession	Source
<i>Chrysomya albiceps</i>	JX913736	Nelson et al. 2012
<i>Chrysomya bezziana</i>	JX913737	Nelson et al., 2012
<i>Chrysomya megacephala1</i>	JX913738	Nelson et al., 2012
<i>Chrysomya megacephala2</i>	JX913739	Nelson et al., 2012
<i>Chrysomya rufifacies</i>	JX913740	Nelson et al., 2012
<i>Chrysomya saffrana</i>	JX913742	Nelson et al., 2012
<i>Chrysomya pinguis</i>	KM244730	Yan et al., 2014
<i>Chrysomya putoria</i>	AF352790	Junqueria et al., 2001
<i>Lucilia sericata</i>	AJ422212	Stevens et al., 2008
<i>Cochliomyia hominivorax</i>	AF260826	Lessinger et al., 2000

The CLC software defines the pairwise parameters as the following.

“Gaps: Calculates the number of alignment positions where one sequence has a gap and the other does not. Differences: Calculates the number of alignment positions where one sequence is different from the other. This includes gap differences as in the Gaps comparison. Distance: Calculates the Jukes-Cantor distance between the two sequences. This number is given as the Jukes-Cantor

correction of the proportion between identical and overlapping alignment positions between the two sequences. Percent identity: Calculates the percentage of identical residues in alignment positions to overlapping alignment positions between the two sequences" (CLC Manual). Figures 4.4.1 and 4.4.2 show the Distance between each listed species is consistent when looking at either COX1 or the entire mitochondrial genome. This same trend extends to Differences as well. The Gap pairwise comparisons, Figures 4.4.3 and 4.4.4 are skewed by the differences in sequence length between the genome and one of its genes. One of the *C. megacephala* assemblies, 43, has a patch of sequence that is of particular low quality, causing eight noncontiguous bases to be missing. The biggest difference between the BOLD sequence and the whole genome data lies within the Percent Identity pairwise comparison. The comparison using only the whole mitochondrial genome shows only two species who share a very high Percent Identity value of 98.74%: Mega2 and *C. saffraneae*. That same value cut-off for the comparison made with only the cytochrome oxidase subunit 1 data includes the first eight listed species. By assessing whole mitochondrial genomic sequencing data, instead of just the COX1 gene, a greater level of distinction can be made when using the Percent Identity tool. To further clarify the added potential of using whole mitochondrial genomic information, Figure 4.4.5a shows a phylogenetic tree made using only the barcode region while Figure 6.4.5b shows one made with the entire genome.

FIGURE 6.4.1: PAIRWISE COMPARISON OF 13 *CHRYSOMYA* FLIES, *LUCILIA SERICATA*, AND *COCHLIOMYIA HOMNIVORAX* USING THE COX1 GENE FOR DIFFERENCES AND DISTANCE

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C. pacifica	1		4	4	4	12	4	2	7	30	68	60	96	111	138	125
42 C. megacephala	2	0.00		0	0	8	0	4	7	32	68	64	94	109	140	127
40 C. megacephala	3	0.00	0.00		0	8	0	4	7	32	68	64	94	109	140	127
41 C. megacephala	4	0.00	0.00	0.00		8	0	4	7	32	68	64	94	109	140	127
43 C. megacephala	5	0.00	0.00	0.00	0.00		8	12	15	40	76	72	102	117	148	135
Mega1	6	0.00	0.00	0.00	0.00	0.00		4	7	32	68	64	94	109	140	127
C. saffranae	7	0.00	0.00	0.00	0.00	0.00	0.00		7	30	70	62	98	113	140	127
Mega2	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00		35	67	61	97	112	137	132
C. pinguis	9	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02		72	67	110	116	152	133
C. bezziana	10	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05		74	109	122	149	140
C. putoria	11	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05		102	113	144	135
C. albiceps	12	0.07	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.08	0.07	0.07		49	153	147
C. ruffacies	13	0.08	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.03		160	152
Cochliomyia hominivorax	14	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.11	0.10	0.10	0.11	0.11		149
Lucilia sericata	15	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.10	0.09	0.10	0.11	0.10	

Top: Differences. Bottom: Distance

FIGURE 6.4.2: PAIRWISE COMPARISON OF 13 *CHRYSOMYA* FLIES, *LUCILIA SERICATA*, AND *COCHLIOMYIA HOMNIVORAX* USING THE MITOCHONDRIAL GENOME FOR DIFFERENCES AND DISTANCE

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C. pacifica	1		830	830	830	839	640	197	197	485	1285	1038	1603	1657	1815	3102
42 C. megacephala	2	0.00		4	2	9	343	954	989	1233	933	1696	1613	1543	2387	2322
40 C. megacephala	3	0.00	0.00		2	13	347	956	991	1235	935	1698	1616	1546	2389	2324
41 C. megacephala	4	0.00	0.00	0.00		11	345	954	989	1233	933	1696	1614	1544	2387	2322
43 C. megacephala	5	0.00	0.00	0.00	0.00		352	963	998	1242	942	1705	1622	1552	2396	2331
Mega1	6	0.00	0.00	0.00	0.00	0.00		622	656	933	707	1445	1444	1432	2198	2665
C. saffranae	7	0.00	0.00	0.00	0.00	0.00	0.00		149	422	1266	968	1614	1680	1758	3221
Mega2	8	0.00	0.01	0.01	0.01	0.01	0.01	0.01		393	1252	960	1594	1651	1760	3219
C. pinguis	9	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02		1294	999	1665	1710	1803	3278
C. bezziana	10	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05		1609	1513	1516	2288	2738
C. putoria	11	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.07		1567	1646	1763	3216
C. albiceps	12	0.08	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08		616	2139	2975
C. ruffacies	13	0.08	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.03		2217	2896
Cochliomyia hominivorax	14	0.10	0.09	0.09	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10		3607
Lucilia sericata	15	0.07	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.08	0.08	0.07	0.08	0.08	0.09	

Top: Differences. Bottom: Distance

FIGURE 6.4.3: PAIRWISE COMPARISON OF 13 *CHRYSOMYA* FLIES, *LUCILIA SERICATA*, AND *COCHLIOMYIA HOMNIVORAX* USING THE COX1 GENE FOR GAPS AND PERCENT IDENTITY

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C. pacifica	1		0	0	0	8	0	0	0	0	0	0	0	0	0	0
42 C. megacephala	2	99.74		0	0	8	0	0	0	0	0	0	0	0	0	0
40 C. megacephala	3	99.74	100.00		0	8	0	0	0	0	0	0	0	0	0	0
41 C. megacephala	4	99.74	100.00	100.00		8	0	0	0	0	0	0	0	0	0	0
43 C. megacephala	5	99.22	99.48	99.48	99.48		8	8	8	8	8	8	8	8	8	8
Mega1	6	99.74	100.00	100.00	100.00	99.48		0	0	0	0	0	0	0	0	0
C. saffranae	7	99.87	99.74	99.74	99.74	99.22	99.74		0	0	0	0	0	0	0	0
Mega2	8	99.54	99.54	99.54	99.54	99.02	99.54	99.54		0	0	0	0	0	0	0
C. pinguis	9	98.04	97.91	97.91	97.91	97.40	97.91	98.04	97.72		0	0	0	0	0	0
C. bezziana	10	95.57	95.57	95.57	95.57	95.06	95.57	95.44	95.63	95.31		0	0	0	0	0
C. putoria	11	96.09	95.83	95.83	95.83	95.32	95.83	95.96	96.02	95.63	95.18		0	0	0	0
C. albiceps	12	93.74	93.87	93.87	93.87	93.37	93.87	93.61	93.68	92.83	92.89	93.35		0	0	0
C. ruffacies	13	92.76	92.89	92.89	92.89	92.39	92.89	92.63	92.70	92.44	92.05	92.63	96.81		0	0
Cochliomyia hominivorax	14	91.00	90.87	90.87	90.87	90.38	90.87	90.87	91.07	90.09	90.29	90.61	90.03	89.57		0
Lucilia sericata	15	91.85	91.72	91.72	91.72	91.22	91.72	91.72	91.40	91.33	90.87	91.20	90.42	90.09	90.29	

Top: Gaps. Bottom: Percent Identity.

FIGURE 6.4.4: PAIRWISE COMPARISON OF 13 *CHRYSOMYA* FLIES, *LUCILIA SERICATA*, AND *COCHLIOMYIA HOMNIVORAX* USING THE MITOCHONDRIAL GENOME FOR GAPS AND PERCENT IDENTITY

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C. pacifica	1		790	789	789	798	593	127	119	134	648	161	467	518	378	2051
42 C. megacephala	2	94.72		3	1	8	341	911	905	908	320	945	621	534	1146	1263
40 C. megacephala	3	94.72	99.97		2	11	344	912	906	909	321	946	624	537	1147	1266
41 C. megacephala	4	94.72	99.99	99.99		9	342	910	904	907	319	944	622	535	1145	1264
43 C. megacephala	5	94.66	99.94	99.91	99.93		349	919	913	916	328	953	629	542	1154	1271
Mega1	6	95.95	97.75	97.73	97.74	97.70		570	564	575	57	614	378	369	819	1604
C. saffranae	7	98.76	93.98	93.97	93.98	93.92	96.07		12	23	625	76	436	505	267	2172
Mega2	8	98.76	93.75	93.74	93.75	93.70	95.86	99.06		15	617	74	424	493	271	2164
C. pinguis	9	96.94	92.22	92.20	92.21	92.16	94.11	97.34	97.52		622	81	431	496	272	2169
C. bezziana	10	91.87	93.88	93.87	93.88	93.82	95.37	92.01	92.10	91.83		653	397	394	854	1563
C. putoria	11	93.45	89.30	89.29	89.30	89.25	90.89	93.90	93.95	93.71	89.86		410	497	269	2166
C. albiceps	12	89.88	89.61	89.59	89.60	89.55	90.73	89.84	89.96	89.52	90.28	90.13		105	627	1836
C. ruffacies	13	89.53	90.01	89.99	90.00	89.95	90.78	89.42	89.60	89.23	90.23	89.63	96.03		716	1757
Cochliomyia hominivorax	14	88.70	85.13	85.12	85.13	85.08	86.31	89.06	89.04	88.78	85.75	89.03	86.69	86.21		2367
Lucilia sericata	15	81.60	85.55	85.54	85.55	85.50	83.76	81.03	81.03	80.69	83.28	81.05	82.12	82.51	78.99	

Top: Gaps. Bottom: Percent Identity.

FIGURE 6.4.5: PHYLOGENETIC TREES CREATED FROM A) COX1 AND B) WHOLE MITOCHONDRIAL GENOMIC SEQUENCE DATA.

Phylogenetic trees created in CLCbio using Neighbor-Joining which allows for varied rates of evolution. Jukes-Cantor algorithm used for nucleotide distance determination which assumes equal base frequencies and substitution rates. Numbers on branches are inferred lengths. Numbers at nodes are bootstraps values.

The differences between using only a section of the mitochondrial genome and using all of it may not be drastic. The more information that can be used in analyses, the more accurate the results will be. The NGS technologies afford research laboratories the ability to use the whole genome rather than just a few parts of it.

5. CONCLUSION

The assembly and annotation of a novel mitochondrial genome for *Chrysomya pacifica* provides another piece to the puzzle for entomologists and as well as phylogeneticists. As more laboratories access NGS platforms, the standard of using only parts of the mitochondrial genome will be replaced by the incorporation of the whole genome into taxonomic studies. There is no longer a need to develop primers for a few mitochondrial and nuclear genes. The both genomes are now easily obtained

The impact of this study is far reaching. The sequencing data and de novo assembly created from four *C. megacephala* and two *C. pacifica* can be used in future work in many fields. The two sister species offer unique insights into how recent evolutionary divergence visible in their phenotypes can affect their genotypes. Much can be learned about how pressures from the environmental niches each inhabit caused their adaptations. One clear example is that *C. megacephala* has a morphologically distinct compound eye that may be linked its very successful adaptation to urban environments. The de novo genome

assembly can provide information on how genes that contribute to compound eye development differ between the two species.

The multitude of uses of genomic data and its newfound place in smaller research laboratories will have a profound effect on the future of science. Draft genomes can be produced in modest studies, and this will cause an explosion of information that was impossible with traditional sequencing methods.

6. REFERENCES

- Alkan, C., Sajjadian, S., & Eichler, E. E. (2011). Limitations of next-generation genome sequence assembly. *Nature Methods*, 8(1), 61-65.
- Anderson, M. W., & Schrijver, I. (2010). Next generation DNA sequencing and the future of genomic medicine. *Genes*, 1(1), 38-69.
- Bao, F., & Wells, J. D. (2014). Population genetic structure of an invasive forensically important insect. *Electrophoresis*, 35(21-22), 3193-3200.
- Bernt, M., Braband, A., Middendorf, M., Misof, B., Rota-Stabelli, O., & Stadler, P. F. (2013). Bioinformatics methods for the comparative analysis of metazoan mitochondrial genome sequences. *Molecular Phylogenetics and Evolution*, 69(2), 320-327.
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritsch, G., et al. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69(2), 313-319.
- Bharti, M., & Kurahashi, H. (2009). Finding of feral derived form (fdf) of *Chrysomya megacephala* (fabricius) from India with an evolutionary novelty (Diptera, Calliphoridae). *Japanese Journal of Systematic Entomology*, 15(2), 411-413.
- Bradnam, K. R., Fass, J. N., Alexandrov, A., Baranay, P., Bechner, M., Birol, et al. (2013). Assemblathon 2: Evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience*, 2(1), 1-31.
- Brower, A. V. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 91(14), 6491-6495.

- Cameron, S. L., Lambkin, C. L., Barker, S. C., & Whiting, M. F. (2007). A mitochondrial genome phylogeny of Diptera: Whole genome sequence data accurately resolve relationships over broad timescales with high precision. *Systematic Entomology*, 32(1), 40-59.
- Cattonaro, F., Policriti, A., & Vezzi, F. (2010). Enhanced reference guided assembly. *Bioinformatics and Biomedicine (BIBM), 2010 IEEE International Conference on*, 77-80.
- Chong, Y. V., Chua, T. H., & Song, B. K. (2014). Genetic variations of *Chrysomya megacephala* populations in Malaysia (Diptera: Calliphoridae). *Advances in Entomology*, 2014
- Cicbio.com. (2012). White paper on de novo assembly in CLC assembly cell 4.0.
- Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics*, 5(6), 435-445.
- El-Metwally, S., Hamza, T., Zakaria, M., & Helmy, M. (2013). Next-generation sequence assembly: Four stages of data processing and computational challenges. *PLoS Computational Biology*, 9(12), e1003345.
- Faircloth, B. C. (2008). Msatcommander: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, 8(1), 92-94.
- Farncombe, K., Beresford, D., & Kyle, C. (2014). Characterization of microsatellite loci in *Phormia regina* towards expanding molecular applications in forensic entomology. *Forensic Science International*, 240, 122-125.
- Florin, A., & Gyllenstrand, N. (2002). Isolation and characterization of polymorphic microsatellite markers in the blowflies *Lucilia illustris* and *Lucilia sericata*. *Molecular Ecology Notes*, 2(2), 113-116.

He, L., Wang, S., Miao, X., Wu, H., & Huang, Y. (2007). Identification of necrophagous fly species using ISSR and SCAR markers. *Forensic Science International*, 168(2), 148-153.

The human genome project: A new reality. (Tue, 24 Mar 2015 15:19:19 GMT). Retrieved from <http://www.sanger.ac.uk/about/history/hgp/>

Junqueira, A. C. M., Lessinger, A. C., Torres, T. T., da Silva, F. R., Vettore, A. L., Arruda, P., & Espin, Ana Maria L Azeredo. (2004). The mitochondrial genome of the blowfly *Chrysomya chloropyga* (Diptera: Calliphoridae). *Gene*, 339, 7-15.

Kurahashi, H. (1982). Probable origin of a synanthropic fly *Chrysomya megacephala*, in New Guinea (Diptera:Calliphoridae). *Biogeography and ecology of New Guinea* (pp. 689-698) Springer.

Lane, R. P. (1975). An investigation into blowfly (Diptera: Calliphoridae) succession on corpses. *Journal of Natural History*, 9(5), 581-588.

Lessinger, A., Martins Junqueira, A., Lemos, T., Kemper, E., Da Silva, F., Vettore, A., . . . L Azeredo-Espin, A. (2000). The mitochondrial genome of the primary screwworm fly *Cochliomyia hominivorax* (Diptera: Calliphoridae). *Insect Molecular Biology*, 9(5), 521-529.

Li, R., Fan, W., Tian, G., Zhu, H., He, L., Cai, J., et al. (2010). The sequence and de novo assembly of the giant panda genome. *Nature*, 463(7279), 311-317.

Life Technologies Website. www.lifetechnologies.com. Accessed 03/30/2015

Llopis, J., McCaffery, J. M., Miyawaki, A., Farquhar, M. G., & Tsien, R. Y. (1998). Measurement of cytosolic, mitochondrial, and golgi pH in single living cells with green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), 6803-6808.

Moller, S., Croning, M. D., & Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics (Oxford, England)*, 17(7), 646-653.

- Nelson, L. A., Lambkin, C. L., Batterham, P., Wallman, J. F., Dowton, M., Whiting, M. F., et al. (2012). Beyond barcoding: A mitochondrial genomics approach to molecular phylogenetics and diagnostics of blowflies (Diptera: Calliphoridae). *Gene*, 511(2), 131-142.
- Nicholls, D. G. (1974). The influence of respiration and ATP hydrolysis on the proton-electrochemical gradient across the inner membrane of rat-liver mitochondria as determined by ion distribution. *European Journal of Biochemistry*, 50(1), 305-315.
- Norris, K. (1965). The bionomics of blow flies. *Annual Review of Entomology*, 10(1), 47-68.
- Pettersson, E., Lundeberg, J., & Ahmadian, A. (2009). Generations of sequencing technologies. *Genomics*, 93(2), 105-111.
- Picard, C., Johnston, J., & Tarone, A. (2012). Genome sizes of forensically relevant Diptera. *Journal of Medical Entomology*, 49(1), 192-197.
- Powell, J. M. (1982). The history of plant use and man's impact on the vegetation. *Biogeography and ecology of New Guinea* (pp. 207-227) Springer.
- Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., et al. (2012). A tale of three next generation sequencing platforms: Comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, 13, 341-2164-13-341. doi:10.1186/1471-2164-13-341 [doi]
- Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The barcode of life data system (<http://www.barcodinglife.org>). *Molecular Ecology Notes*, 7(3), 355-364.
- Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., et al. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475(7356), 348-352.

- Rozen, S., & Skaletsky, H. (1999). Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics methods and protocols* (pp. 365-386) Springer.
- Russell, P. (Ed.). (2006). *IGenetics, A mendelian approach* Benjamin-Cummings Publishing Company.
- Sheffield, N. C., Song, H., Cameron, S. L., & Whiting, M. F. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Molecular Biology and Evolution*, 25(11), 2499-2509. doi:10.1093/molbev/msn198 [doi]
- Sims, D. (2014). Sequencing depth and coverage: Key considerations in genomic analyses. *Nature Reviews.Genetics*, 15(2), 121; 121-132; 132.
- Singh, B., & Wells, J. D. (2011). Chrysomyinae (Diptera: Calliphoridae) is monophyletic: A molecular systematic analysis. *Systematic Entomology*, 36(3), 415.
- Slomovic, S., Portnoy, V., Liveanu, V., & Schuster, G. (2006). RNA polyadenylation in prokaryotes and organelles; different tails tell different tales. *Critical Reviews in Plant Sciences*, 25(1), 65; 65-77; 77.
- Smith, C., Heyne, S., Richter, A. S., Will, S., & Backofen, R. (2010). Freiburg RNA tools: A web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Research*, 38(Web Server issue), W373-7. doi:10.1093/nar/gkq316 [doi]
- Stevens, J., West, H., & Wall, R. (2008). Mitochondrial genomes of the sheep blowfly, *Lucilia sericata*, and the secondary blowfly, *Chrysomya megacephala*. *Medical and Veterinary Entomology*, 22(1), 89-91.
- Torres, T. T., AZEREDO-ESPIN, D., & LIMA, A. M. (2008). Characterization of polymorphic microsatellite markers for the

blowfly *Chrysomya albiceps* (Diptera: Calliphoridae). *Molecular Ecology Resources*, 8(1), 208-210.

VÁZQUEZ, J. F., PÉREZ, T., ALBORNOZ, J., & DOMÍNGUEZ, A. (2000). Estimation of microsatellite mutation rates in *Drosophila melanogaster*. *Genetical Research*, 76(03), 323-326.

Wells, J.D. & H. Kurahashi. 1996. A new species of *Chrysomya* (Diptera: Calliphoridae) from Sulawesi, Indonesia, with a key to the Oriental, Australasian and Oceanian species. *Medical Entomology and Zoology* 47:131-138, and Whitworth T. 2006. Key to the genera and species of blow flies (Diptera: Calliphoridae) of America north of Mexico. *Proceedings of the Entomological Society of Washington* 108:689-725. Wells, J., & Sperling, F. A. Mitochondrial DNA divergence between *Chrysomya* species (Diptera: Calliphoridae): Recent origin of an urban pest. *Unpublished*,

Wells, J., & Stevens, J. (2008). Applications of DNA-based methods in forensic entomology. *Annals Review of Entomology*, 53, 103-120.

Wells, J. D. (1991). *Chrysomya megacephala* (Diptera: Calliphoridae) has reached the continental united states: Review of its biology, pest status, and spread around the world. *Journal of Medical Entomology*, 28(3), 471-473.

Wells, J. D., Wall, R., & Stevens, J. R. (2007). Phylogenetic analysis of forensically important *Lucilia* flies based on cytochrome oxidase I sequence: A cautionary tale for forensic species determination. *International Journal of Legal Medicine*, 121(3), 229-233.

Wetterstrand, K. (2014). DNA sequencing costs: Data from the NHGRI genome sequencing program (GSP). Retrieved from <http://www.genome.gov/sequencingcosts>

Yan, J., Liao, H., Xie, K., & Cai, J. (2014). The complete mitochondria genome of *Chrysomya pinguis* (Diptera: Calliphoridae). *Mitochondrial DNA*, (0), 1-3

APPENDIX A
AGAROSE GEL ELECTROPHORESIS RESULTS

40-10017 STR:
 F:GAAGAGCAGAACGAGAAA
 R:ACCTACTACTATCTCTACC
 EST: 93b (AAAT)n
 Tm: 49.7
 40,41,43 :290b

40-39693 ISSR:
 F:CCAACACAAACAACAACAAC
 R: CTGTTGTGGTTGTTGTTGT
 EST: 234B AAC
 Tm: 50.5

40-15840 STR:
 F:CGACTAAAACCTACTTCAAC
 R: CACCCCATCAATTACAAC
 EST: 97b AAC
 Tm: 48.2
 38Pac:82
 43Mega : 94
 40,41,42 : 97b

40-39693 STR:
 F:TGATGGTGGTGTGTTTTG
 R:AGAAATAAACAGGGCGCA
 EST:290b
 Tm:46
 38:214
 40:239
 41:217
 43:213

38-7057 ISSR:
 F:GTTGTTGTTGTTGTTGTCC
 R:FAACAACAACAACAACCG
 EST:423b
 Tm: 50.3

38-28185 ISSR:
 F:CAAGAACAACAACAACAACC
 R:GTTGTTGTTGTTGTTGTTGC
 EST:115b
 Tm:50.8

10-10017STR		
78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

40-39693ISSR

78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

40-15840STR

78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

40-39693STR

78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

38-7057ISSR

78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

38-28185ISSR		
78-82	83-87	88-92

93-97	98-102	103-107

108-112	113-117	118-122

123-127	128-132	133-137

APPENDIX B

CHRYSOMYA PACIFICA MITOCHONDRIAL GENOME

ND1

C. pacifica AAGTTCCACGTA CTCAAAATAAATATAAATGAACAAATGTTAACTTAATATAAAAAAGAAAGAAATACATCTCTTCCTAAGAATATTACACAAAATAATATACTTATAAAATAAAATAC

ND1

C. pacifica TTGCATATTCAGCCAAAAAATTAATGCAAATCCTCCTCTTCTATATTCTACATTAAATCCAGAACTAATTCTGATTCTCTTCCGCAAAATCAAAGGTGTACGATTTGTTTCTGCTA

ND1

C. pacifica AAGAAATACTAAATCAAACCTAAAGCTATAGGAAATATAAATAAAAAATCAAATAAATATTTGATACTTATAAAATATTAAATATTATAACCCCAATTAAAAAAATAAACTTAATA

ND1

C. pacifica AAATAATGCTAAACTTACTTCATAAGAAATAGTTTGAGCAACAGCTCGTAACCTCCTAATAAAGCATAATTAGAATTAGAAGATCAACCAGCAATTATTACTGTATAAACCCCAAC

ND1

C. pacifica TAGTACAACATAAGAAAAACAATAAACCTAAATTAAAAGAAAAAGCTTTACAATATTGGTATACATATTCATACTAATAAAGACAAAAATAAGAAAAAATAGGTGAAAAATAATAAG

ND1

C. pacifica AAATATAATTAGACAACAAAGGATAAGTTTGTTCTTAGTAAATAACTTAATTGCATCACAAAAAGTTGAGGAATTCCTGCAATTCCAACCTTATTAGGACCTTTTGAATCTGAATAT

ND1

C. pacifica ATCCTAATACCTTTCGTTCTAAAAGAGTTAAAAAGCTACTCTTACTAATACAAAAATAATTAATAATCTTCCAATAATAAATAAAATATTTTCTATAAAAAACAAGTACTATTTGT

trnaL (tag)

C. pacifica GATAAAATCACATACATAAATTCTAAATTTATTGCACTAATCTGCCAAATAGTTATTTATATTAAATTATTTCAATATAAAAAATAAATTTATCAAAATATTAGGTCCTTTTCGTA CTG

trnaL (tag)

C. pacifica AAATATTTTAATTTTTTAAAGATAGAAACCAACCTGCTTACGCCGTTTGAACCTCAGATCATGTAAGAATTTAAAGTCGAACAGACTTAAATTTAAGCGGCTACACCTAAAATTATAT

trnaL

C. pacifica CTTAATCCAACATCGAGGTCGCAATCTTTTTATCGATATGAACTCTCCAAAAAATTACGCTGTTATCCCTAAAGTAACCTATTTTTTAATCGCTAATAACGGATCAATTATTCATAA

trnaL

C. pacifica ATTAATGATTTAAACAATTAAGTTTATCAAAATTTAATATCACCCCAATAAAATATTATCAATTATTATAACAAAAAATCTATAAAATTATAATAACCTAAATATAAGATTATTA

trnaL

C. pacifica GGGTCTTCTCGTCTTTTAAATATATTTTAGCTTTTTGACTAAAAATAAAATCTATTATAAATTTTTATGAAACAGTTAATATCTCGTCCAACCATTCATACCAGCCTTCAATTAAAG

trnaL

C. pacifica ACTAATGATTATGCTACCTTTGCACAGTTAGTATACTGCGGCCATTTAAAAATTATTCAGTGGGCAGGATAGACTTTAAAAAATTCAAAAAGACATGTTTTTGTAAACAGGCGAACA

trnaL

C. pacifica TTATTTTTGCCGAGTTCTTTATAAAAACCTTTTCATTTATATTTATTTATACAATACAATATACTAATTTTATCATTATTTATTTATTTTATAAATTAATAAATATTTCAATAAAACT

trnaL

C. pacifica TAAATTTTAATAAATAAATAATTATTATAAAATAAATTATAACACTTTTATTAATAATAGCTATTTCTAAGCTTATTTTATTAACCTATTTTATTTTATTTTAAAAATTTATTTTACAGC

	15,500	15,520	15,540	15,560	15,580	15,600
<i>C. pacifica</i>	TTATATAAATAATTTATTAATTATTAATTTTATAACATTATATTCATAATTTAATATATTATAAATATATGAATAAATATATATAAATTATATATATATATATATATGCAATTTTAAATT					
	15,620	15,640	15,660	15,680	15,700	
<i>C. pacifica</i>	ATCTAAATTTATTTTTTTTCACTTTATAAACTAATTTACTATTTATAATTATTAATATTTAAATTTTATAATATTATTTTATTTAATAAATCACAAAAAANNACTATAAAATAAA					

APPENDIX C

ALIGNMENT OF *C. PACIFICA* AND *C. MEGACEPHALA* GENOMES

ND2	320	340	360	380	400																													
C. pacifica	GGAGCTTGAA	TAGGTTTAGA	AATTAATTTA	TTAGCTTTTA	TCCCCCTAAT	AAGAGATAAT	AAATTAATAT	CAACAGAAGC	CTCATTAAAG	TATTTTCTTA	400																							
Translation ORF/CDS	G	A	W	M	G	L	E	I	N	L	L	A	F	I	P	L	M	S	D	N	K	L	M	S	T	E	A	S	L	K	Y	F	L	
ND2																																		
Mega1																																		
Translation ORF/CDS	G	A	W	M	G	L	E	I	N	L	L	A	F	I	P	L	M	S	D	N	K	L	M	S	T	E	A	S	L	K	Y	F	L	
ND2																																		
Mega2																																	C	
Translation ORF/CDS	G	A	W	M	G	L	E	I	N	L	L	A	F	I	P	L	M	S	D	N	K	L	M	S	T	E	A	S	L	K	Y	F	L	
Consensus	GGAGCTTGAA	TAGGTTTAGA	AATTAATTTA	TTAGCTTTTA	TCCCCCTAAT	AAGAGATAAT	AAATTAATAT	CAACAGAAGC	CTCATTAAAG	TATTT	CTTA																							
Conservation																																		
ND2	420	440	460	480	500																													
C. pacifica	CTCAAGCATT	AGCTTCTTCA	GTATTCCTAT	TTGCAGTAAT	TTTATTTT	CTAAATTCAA	GTAAGAGAAA	TTCAAATTAC	TTTATAGAAA	TAATGATTTT	500																							
Translation ORF/CDS	T	Q	A	L	A	S	S	V	F	L	F	A	V	I	L	F	L	L	N	S	S	K	S	N	S	N	Y	F	M	E	M	M	I	F
ND2																																		
Mega1																																		
Translation ORF/CDS	T	Q	A	L	A	S	S	V	F	L	F	A	V	I	L	F	L	L	N	S	S	K	S	N	S	N	Y	F	M	E	M	M	I	F
ND2																																		
Mega2																																		
Translation ORF/CDS	T	Q	A	L	A	S	S	V	F	L	F	A	V	I	L	F	L	L	N	S	S	K	S	N	S	N	Y	F	M	E	M	M	I	F
Consensus	CTCAAGCATT	AGCTTCTTCA	GTATTCCTAT	TTGCAGTAAT	TTTATTTT	CTAAATTCAA	GTAAGAGAAA	TTCAAATTAC	TTTATAGAAA	TAATGATTTT																								
Conservation																																		
ND2	520	540	560	580	600																													
C. pacifica	CTCTTCTCTT	TTATTAAAA	GAGGATCAGC	TCCTTTCCAC	TTTTGATTTC	CTAATGTAAT	GGAAGGACTA	TCTTGATTAA	ATGCATTAAT	TCTAATAACT	600																							
Translation ORF/CDS	S	S	L	L	L	K	S	G	S	A	P	F	H	F	W	F	P	N	V	M	E	G	L	S	W	L	N	A	L	I	L	M	T	
ND2																																		
Mega1																																		
Translation ORF/CDS	S	S	L	L	L	K	S	G	S	A	P	F	H	F	W	F	P	N	V	M	E	G	L	S	W	L	N	A	L	I	L	M	T	
ND2																																		
Mega2																																		
Translation ORF/CDS	S	S	L	L	L	K	S	G	S	A	P	F	H	F	W	F	P	N	V	M	E	G	L	S	W	L	N	A	L	I	L	M	T	
Consensus	CTCTTCTCTT	TTATTAAA	A	GAGGATCAGC	TCCTTTCCAC	TTTTGATT	C	CTAATGTAAT	GGAAGGACTA	TCTTGATTAA	ATGCATTAAT	TCTAATAACT																						
Conservation																																		

	ND2	620	640	660	680	700																											
C. pacifica	TGACAAAAAA	TTGCTCCTTT	AATATTAATT	TCTTATATTA	TCTTTAAGCC	ATTAATTATT	ACAAGCATT	TTCTTTCTAG	TTTAATTGGT	GCATTAGGGG	700																						
Translation ORF/CDS	W	Q	K	A	P	L	M	L	S	Y	F	K	P	L	T	S	L	S	S	L	G	A	L	G									
	ND2																																
	Mega1																																
Translation ORF/CDS	W	Q	K	A	P	L	M	L	S	Y	F	K	P	L	T	S	L	S	S	L	G	A	L	G									
	ND2																																
	Mega2																																
Translation ORF/CDS	W	Q	K	A	P	L	M	L	S	Y	F	K	P	L	T	S	L	S	S	L	G	A	L	G									
Consensus	TGACAAAAAA	TTGCTCCTTT	AATATTAATT	TCTTATATTA	TCTTTAAGCC	ATTAATTATT	ACAAGCATT	TTCTTTCTAG	TTTAATTGGT	GCATTAGGGG																							
Conservation																																	
	ND2	720	740	760	780	800																											
C. pacifica	GATTAATCA	AACTTCACTA	CGTAAATTAA	TAGCTTATTC	TTCAATTAAT	CATCTTGGAT	GAATATTAGC	AGCAATATAC	AATAGTAATT	TACTATGAAT	800																						
Translation ORF/CDS	G	L	N	Q	T	S	L	R	K	L	M	A	Y	S	S	N	H	L	G	W	M	L	A	A	M	Y	N	S	N	L	L	W	M
	ND2																																
	Mega1																																
Translation ORF/CDS	G	L	N	Q	T	S	L	R	K	L	M	A	Y	S	S	N	H	L	G	W	M	L	A	A	M	Y	N	S	N	L	L	W	M
	ND2																																
	Mega2																																
Translation ORF/CDS	G	L	N	Q	T	S	L	R	K	L	M	A	Y	S	S	N	H	L	G	W	M	L	A	A	M	Y	N	S	N	L	L	W	M
Consensus	GATTAATCA	AACTTCACTA	CGTAAATTAA	TAGCTTATTC	TTCAATTAAT	CATCTTGGAT	GAATATTAGC	AGCAATATAC	AATAGTAATT	TACTATGAAT																							
Conservation																																	
	ND2	820	840	860	880	900																											
C. pacifica	AACATATTTT	ATATTTTATT	CATTTTAAAC	TTTTAGTATA	ATTTTATAT	TTAATATATT	CAAACTTCC	CATATTAATC	AATTATTTAC	CTTATTTTTT	900																						
Translation ORF/CDS	T	Y	F	M	F	Y	S	F	L	T	F	S	M	I	F	M	F	N	M	F	K	T	S	H	I	N	Q	L	F	T	L	F	F
	ND2																																
	Mega1																																
Translation ORF/CDS	T	Y	F	M	F	Y	S	F	L	T	F	S	M	I	F	M	F	N	M	F	K	T	S	H	I	N	Q	L	F	T	L	F	F
	ND2																																
	Mega2																																
Translation ORF/CDS	T	Y	F	M	F	Y	S	F	L	T	F	S	M	I	F	M	F	N	M	F	K	T	S	H	I	N	Q	L	F	T	L	F	F
Consensus	AACATATTTT	ATATTTTATT	CATTTTAAAC	TTTTAGTATA	ATTTTATAT	TTAATATATT	CAAACTTCC	CATATTAATC	AATTATTTAC	CTTATTTTTT																							
Conservation																																	

	ND2	920	940	960	980	1,000	
C. pacifica	CATTCAAAGG TTATAAAGTT TTTCTTATTC TTAAATTTAT TATCCTTAGG TGGATTACCT CCATTTTITAG GATTTTTCCT AAAATGAATT GTTATTCAAT						1000
Translation ORF/CDS	H S K V M K F F L F F N L L S L G G L P P F L G F F P K W V Q						
	ND2						
Mega1							1000
Translation ORF/CDS	H S K V M K F F L F F N L L S L G G L P P F L G F F P K W V Q						
	ND2						
Mega2							1000
Translation ORF/CDS	H S K V M K F F L F F N L L S L G G L P P F L G F F P K W V Q						
Consensus	CATTCAAAGG TTATAAAGTT TTTCTTATTC TTAAATTTAT TATCCTTAGG TGGATTACCT CCATTTTITAG GATTTTTCCT AAAATGAATT GTTATTCAAT						
Conservation							
	ND2	1,020	1,040	1,060	1,080	1,100	
C. pacifica	CATTAACATAT AAATAACCAG TTATTTTITAT TAACATTTAT GGTTTAAATG ACTTTAATTA CTTTATACCT TTATATACGA TTATGTTATA GAGCATTAT						1100
Translation ORF/CDS	S L T M N N Q L F L L T F M V L M T L I T L Y F Y M R L C Y S A F M						
	ND2						
Mega1							1100
Translation ORF/CDS	S L T M N N Q L F L L T F M V L M T L I T L Y F Y M R L C Y S A F M						
	ND2						
Mega2							1100
Translation ORF/CDS	S L T M N N Q L F L L T F M V L M T L I T L Y F Y M R L C Y S A F M						
Consensus	CATTAACATAT AAATAACCAG TTATTTTITAT TAACATTTAT GGTTTAAATG ACTTTAATTA CTTTATACCT TTATATACGA TTATGTTATA GAGCATTAT						
Conservation							
	ND2	1,120	1,140	1,160	1,180	1,200	
C. pacifica	ATTAAATTAT TATGAAAATA ACTGATTAAC TGCATCTATC TATAAACTT CTTATATAAA GGTATTATTA ATTTGTTTAT TTTTCTCTC ATTTGGGTTA						1200
Translation ORF/CDS	L N Y Y E N N W L T A S I Y K T S Y M K V F M I C S F F S S F G L						
	ND2						
Mega1							1200
Translation ORF/CDS	L N Y Y E N N W L T A S I Y K T S Y M K V F M I C S F F S S F G L						
	ND2						
Mega2							1200
Translation ORF/CDS	L N Y Y E N N W L T A S I Y K T S Y M K V F M I C S F F S S F G L						
Consensus	ATTAAATTAT TATGAAAATA ACTGATTAAC TGCATCTATC TATAAACTT CTTATATAAA GGTATTATTA ATTTGTTTAT TTTTCTCTC ATTTGGGTTA						
Conservation

COX1	1,520	1,540	1,560	1,580	1,600																												
C. pacifica	GGAAATAGTAG	GAAC TTCATT	AAGTATTTTA	ATT CGAGCTG	AATTAGGACA	CCCTGGAGCA	TTAATTGGAG	ACGACCAAAT	TTATAATGTA	ATTGTAACAG	1599																						
Translation ORF/CDS	G	M	V	G	T	S	L	S	L	R	A	E	L	G	H	P	G	A	L	G	D	D	Q	Y	N	V	V	T					
COX1																																	
Mega1																																	
Translation ORF/CDS	G	M	V	G	T	S	L	S	L	R	A	E	L	G	H	P	G	A	L	G	D	D	Q	Y	N	V	V	T					
COX1																																	
Mega2																																	
Translation ORF/CDS	G	M	V	G	T	S	L	S	L	R	A	E	L	G	H	P	G	A	L	G	D	D	Q	Y	N	V	V	T					
Consensus	GGAAATAGTAG	GAAC TTCATT	AAGTATTTTA	ATT CGAGCTG	AATTAGGACA	CCCTGGAGCA	TTAATTGGAG	ACGACCAAAT	TTATAATGTA	ATTGTAACAG																							
Conservation																																	
COX1	1,620	1,640	1,660	1,680	1,700																												
C. pacifica	CTCACGCTTT	TATTATAATT	TTCTTTATAG	TAATGCCAAT	TATAATTGGA	GGATTTGGAA	ATTGACTAGT	TCCTTTAATG	TTAGGAGCTC	CAGATATAGC	1699																						
Translation ORF/CDS	A	H	A	F	M	F	F	M	V	M	P	M	G	G	F	G	N	W	L	V	P	L	M	L	G	A	P	D	M	A			
COX1																																	
Mega1																																	
Translation ORF/CDS	A	H	A	F	M	F	F	M	V	M	P	M	G	G	F	G	N	W	L	V	P	L	M	L	G	A	P	D	M	A			
COX1																																	
Mega2																																	
Translation ORF/CDS	A	H	A	F	M	F	F	M	V	M	P	M	G	G	F	G	N	W	L	V	P	L	M	L	G	A	P	D	M	A			
Consensus	CTCACGCTTT	TATTATAATT	TTCTTTATAG	TAATGCCAAT	TATAATTGGA	GGATTTGGAA	ATTGACTAGT	TCCTTTAATG	TTAGGAGCTC	CAGATATAGC																							
Conservation																																	
COX1	1,720	1,740	1,760	1,780	1,800																												
C. pacifica	TTTCCCACGA	ATAAATAATA	TAAGTTTCTG	ACTTTTACCT	CCTGCATTAA	CTTTATTATT	AGTAAGTAGT	ATAGTAGAAA	ATGGAGCTGG	AACAGGATGA	1799																						
Translation ORF/CDS	F	P	R	M	N	N	M	S	F	W	L	L	P	P	A	L	T	L	L	L	V	S	S	M	V	E	N	G	A	G	T	G	W
COX1																																	
Mega1																																	
Translation ORF/CDS	F	P	R	M	N	N	M	S	F	W	L	L	P	P	A	L	T	L	L	L	V	S	S	M	V	E	N	G	A	G	T	G	W
COX1																																	
Mega2																																	
Translation ORF/CDS	F	P	R	M	N	N	M	S	F	W	L	L	P	P	A	L	T	L	L	L	V	S	S	M	V	E	N	G	A	G	T	G	W
Consensus	TTTCCCACGA	ATAAATAATA	TAAGTTTCTG	ACTTTTACCT	CCTGCATTAA	CTTTATTATT	AGTAAGTAGT	ATAGTAGAAA	ATGGAGCTGG	AACAGGATGA																							
Conservation																																	

	COX1	1,820	1,840	1,860	1,880	1,900	
C. pacifica	ACTGTTTACC CACCTTTATC TTCTAATATT GCTCATGGAG GAGCATCAGT TGATTTAGCT ATTTTCTCTT TACACTTAGC AGGAATTTCT TCAATTTTAG						1899
Translation ORF/CDS	T V Y P P L S S N A H G G A S V D L A F S L H L A G S S L						
	COX1						
Mega1							1899
Translation ORF/CDS	T V Y P P L S S N A H G G A S V D L A F S L H L A G S S L						
	COX1						
Mega2							1900
Translation ORF/CDS	T V Y P P L S S N A H G G A S V D L A F S L H L A G S S L						
Consensus	ACTGTTTACC CACCTTTATC TTCTAATATT GCTCATGGAG GAGCATCAGT TGATTTAGCT ATTTTCTCTT TACACTTAGC AGGAATTTCT TCAATTTTAG						
Conservation							
	COX1	1,920	1,940	1,960	1,980	2,000	
C. pacifica	GAGCTGTAAG TTTTATTACA ACTGTAATTA ATATACGATC TACAGGAATT ACATTGATC GAATACCTTT ATTTGTATGA TCTGTAGTTA TTACTGCTCT						1999
Translation ORF/CDS	G A V N F T T V N M R S T G T F D R M P L F V W S V V T A L						
	COX1						
Mega1							1999
Translation ORF/CDS	G A V N F T T V N M R S T G T F D R M P L F V W S V V T A L						
	COX1						
Mega2							2000
Translation ORF/CDS	G A V N F T T V N M R S T G T F D R M P L F V W S V V T A L						
Consensus	GAGCTGTAAG TTTTATTACA ACTGTAATTA ATATACGATC TACAGGAATT ACATTGATC GAATACCTTT ATTTGTATGA TCTGTAGTTA TTACTGCTCT						
Conservation							
	COX1	2,020	2,040	2,060	2,080	2,100	
C. pacifica	ATTATTATTA TTATCTTTAC CAGTATTAGC TGGAGCTATT ACTATATTAT TAACTGACCG AAATCTAAAT ACTTCATTCT TTGACCCAGC AGGAGGAGGA						2099
Translation ORF/CDS	L L L L S L P V L A G A T M L L T D R N L N T S F F D P A G G G						
	COX1						
Mega1							2099
Translation ORF/CDS	L L L L S L P V L A G A T M L L T D R N L N T S F F D P A G G G						
	COX1						
Mega2							2100
Translation ORF/CDS	L L L L S L P V L A G A T M L L T D R N L N T S F F D P A G G G						
Consensus	ATTATTATTA TTATCTTTAC CAGTATTAGC TGGAGCTATT ACTATATTAT TAACTGACCG AAATCTAAAT ACTTCATTCT TTGACCCAGC AGGAGGAGGA						
Conservation							

COX1	2,120	2,140	2,160	2,180	2,200
C. pacifica	GATCCTATTT TATACCAACA TTTATTTTGA TTCTTTGGAC ATCCTGAAGT TTATATTTTA ATTTTACCTG GATTCCGGAAT AATTTCTCAT ATTATTAGTC				
Translation ORF/CDS	D P L Y Q H L F W F F G H P E V Y L L P G F G M S H S				
COX1					
Mega1	T.....				
Translation ORF/CDS	D P L Y Q H L F W F F G H P E V Y L L P G F G M S H S				
COX1					
Mega2					
Translation ORF/CDS	D P L Y Q H L F W F F G H P E V Y L L P G F G M S H S				
Consensus	GATCCTATTT TATA-CAACA TTTATTTTGA TTCTTTGGAC ATCCTGAAGT TTATATTTTA ATTTTACCTG GATTCCGGAAT AATTTCTCAT ATTATTAGTC				
Conservation					
COX1	2,320	2,340	2,360	2,380	2,399
C. pacifica	AAGAATCAGG AAAGAAGGAA ACTTTCGGAT CTTTAGGAAT GATTTATGCT ATACTAGCTA TTGGTCTATT AGGATTTATT GTATGAGCTC ACCACATGTT				
Translation ORF/CDS	Q E S G K K E T F G S L G M Y A M L A G L L G F V W A H H M F				
COX1					
Mega1	A.....				
Translation ORF/CDS	Q E S G K K E T F G S L G M Y A M L A G L L G F V W A H H M F				
COX1					
Mega2	A.....				
Translation ORF/CDS	Q E S G K K E T F G S L G M Y A M L A G L L G F V W A H H M F				
Consensus	AAGAATCAGG AAA-AAGGAA ACTTTCGGAT CTTTAGGAAT -ATTTATGCT ATACTAGCTA TTGGTCTATT AGGATTTATT GTATGAGCTC ACCACATGTT				
Conservation					
COX1	2,320	2,340	2,360	2,380	2,400
C. pacifica	TACTGTTGGA ATAGACGTAG ACACACGAGC TTATTTCACT TCAGCTACAA TAATTATTGC TGTACCAACT GGAATTAAGA TTTTCAGTTG ATTAGCAACT				
Translation ORF/CDS	T V G M D V D T R A Y F T S A T M A V P T G K F S W L A T				
COX1					
Mega1					
Translation ORF/CDS	T V G M D V D T R A Y F T S A T M A V P T G K F S W L A T				
COX1					
Mega2					
Translation ORF/CDS	T V G M D V D T R A Y F T S A T M A V P T G K F S W L A T				
Consensus	TACTGTTGGA ATAGACGTAG ACACACGAGC TTATTTCACT TCAGCTACAA TAATTATTGC TGTACCAACT GGAATTAAGA TTTTCAGTTG ATTAGCAACT				
Conservation					

	COX1	2,420	2,440	2,460	2,480	2,500	
C. pacifica	CTTTACGGAA CACAATTAAA TTATTCTCCA GCTACTTTAT GAGCTTTAGG ATTTGTATTT TTATTTACTG TAGGAGGATT AACTGGAGTT GTTTTAGCTA						2499
Translation ORF/CDS	L Y G T Q L N Y S P A T L W A L G F V F L F T V G G L T G V V L A						
	COX1						
Mega1							2499
Translation ORF/CDS	L Y G T Q L N Y S P A T L W A L G F V F L F T V G G L T G V V L A						
	COX1						
Mega2							2500
Translation ORF/CDS	L Y G T Q L N Y S P A T L W A L G F V F L F T V G G L T G V V L A						
Consensus	CTTTACGGAA CACAATTAAA TTATTCTCCA GCTACTTTAT GAGCTTTAGG ATTTGTATTT TTATTTACTG TAGGAGGATT AACTGGAGTT GTTTTAGCTA						
Conservation							
	COX1	2,520	2,540	2,560	2,580	2,600	
C. pacifica	ATTTCATCAAT TGACATTATT TTACATGATA CATATTATGT AGTAGCTCAC TTCCATTATG TTCTATCAAT GGGAGCTGTA TTTGCTATTA TAGCAGGATT						2599
Translation ORF/CDS	N S S D L H D T Y Y V V A H F H Y V L S M G A V F A M A G F						
	COX1						
Mega1							2599
Translation ORF/CDS	N S S D L H D T Y Y V V A H F H Y V L S M G A V F A M A G F						
	COX1						
Mega2							2600
Translation ORF/CDS	N S S D L H D T Y Y V V A H F H Y V L S M G A V F A M A G F						
Consensus	ATTTCATCAAT TGACATTATT TTACATGATA CATATTATGT AGTAGCTCAC TTCCATTATG TTCTATCAAT GGGAGCTGTA TTTGCTATTA TAGCAGGATT						
Conservation							
	COX1	2,620	2,640	2,660	2,680	2,700	
C. pacifica	TGTTTCATTGA TTCCCTCTAT TTAAGTGGATT AACTTTAAAT AGCAAGTTAT TAAAGAGTCA ATTTGCTATT ATATTTATCG GAGTAAATTT AACATTCTTC						2699
Translation ORF/CDS	V H W F P L F T G L T L N S K L L K S Q F A M F G V N L T F F						
	COX1						
Mega1							2699
Translation ORF/CDS	V H W F P L F T G L T L N S K L L K S Q F A M F G V N L T F F						
	COX1						
Mega2							2700
Translation ORF/CDS	V H W F P L F T G L T L N G K L L K S Q F A M F G V N L T F F						
Consensus	TGTTTCATTGA TTCCCTCTAT TTAAGTGGATT AACTTTAAAT -GCAA-TTAT TAAAGAGTCA ATTTGCTATT ATATTTATCG GAGTAAATTT AACATTCTTC						
Conservation							

	COX1	2,720	2,740	2,760	2,780	2,800	
C. pacifica	CCTCAACATT TCTTAGGATT AGCAGGTATA CCTCGACGAT ACTCAGACTA TCCAGACGCT TACACAGCTT GAAATGTAAT TTCTACAATT GGTTCACAA						2799
Translation ORF/CDS	P Q H F L G L A G M P R R Y S D Y P D A Y T A W N V S T G S T						
	COX1						
Mega1							2799
Translation ORF/CDS	P Q H F L G L A G M P R R Y S D Y P D A Y T A W N V S T G S T						
	COX1						
Mega2							2800
Translation ORF/CDS	P Q H F L G L A G M P R R Y S D Y P D A Y T A W N V S T G S T						
Consensus	CCTCAACATT TCTTAGGATT AGCAGGTATA CCTCGACGAT ACTCAGACTA TCCAGACGCT TACACAGCTT GAAATGTAAT TTCTACAATT GGTTCACAA						
Conservation							
	COX1	2,820	2,840	2,860	2,880	2,900	
C. pacifica	TTTCATTATT AGGAATTTTA TTCTTCTTTT TCATTATTTG AGAAAGTTTA GTATCTCAAC GACGAGTTTT ATTCCCTGTT CAACTAAATT CATCAATTGA						2899
Translation ORF/CDS	S L L G L F F F F W E S L V S Q R R V L F P V Q L N S S E						
	COX1						
Mega1							2899
Translation ORF/CDS	S L L G L F F F F W E S L V S Q R R V L F P V Q L N S S E						
	COX1						
Mega2							2900
Translation ORF/CDS	S L L G L F F F F W E S L V S Q R R V L F P V Q L N S S E						
Consensus	TTTCATTATT AGGAATTTTA TTCTTCTTTT TCATTATTTG AGAAAGTTTA GTATCTCAAC GACGAGTTTT ATTCCCTGTT CAACTAAATT CATCAATTGA						
Conservation							
	COX1	2,920	2,940	2,960	2,980	3,000	
C. pacifica	ATGATTACAA AATACTCCAC CAGCTGAACA CAGCTATAGT GAATTACCTT TATTAACATA CTTCTAATAT GGCAGATTAG TGCAATGGAT TTAAGCTCCA						2999
Translation ORF/CDS	W L Q N T P P A E H S Y S E L P L L T N						
	COX1						
Mega1							2999
Translation ORF/CDS	W L Q N T P P A E H S Y S E L P L L T N						
	COX1						
Mega2							3000
Translation ORF/CDS	W L Q N T P P A E H S Y S E L P L L T N						
Consensus	ATGATTACAA AATACTCCAC CAGCTGAACA CAGCTATAGT GAATTACCTT TATTAACATA CTTCTAATAT GGCAGATTAG TGCAATGGAT TTAAGCTCCA						
Conservation

	COX2	3,320	3,340	3,360	3,380	3,400	
C. pacifica	ATCAATGAAC CTTCAATTAC TTAAAGGCT ATTGGACATC AATGATATTG AAGTTATGAA TATTCAGATT TTGCTAATAT TGAATTTGAT TCATACATAA						3399
Translation ORF/CDS	N E P S T L K A G H Q W Y W S Y E Y S D F A N E F D S Y M						
	COX2						
Mega1							3399
Translation ORF/CDS	N E P S T L K A G H Q W Y W S Y E Y S D F A N E F D S Y M						
	COX2						
Mega2							3400
Translation ORF/CDS	N E P S T L K A G H Q W Y W S Y E Y S D F A N E F D S Y M						
Consensus	ATCAATGAAC CTTCAATTAC TTAAAGGCT ATTGGACATC AATGATATTG AAGTTATGAA TATTCAGATT TTGCTAATAT TGAATTTGAT TCATACATAA						
Conservation							
	COX2	3,420	3,440	3,460	3,480	3,500	
C. pacifica	TTCTACTAA CGAACTATCA ATTGATAGCT TCCGTCTACT AGACGTTGAT AATCGAGTAG TATTACCAAT AAATTCACAA ATTCGAATTT TAGTAACAGC						3499
Translation ORF/CDS	P T N E L S D S F R L L D V D N R V V L P M N S Q R L V T A						
	COX2						
Mega1							3499
Translation ORF/CDS	P T N E L S D S F R L L D V D N R V V L P M N S Q R L V T A						
	COX2						
Mega2	T.....						3500
Translation ORF/CDS	P T N E L S D S F R L L D V D N R V V L P M N S Q R L V T A						
Consensus	TTCTACTAA CGAACTATCA ATTGATAG-T TCCGTCTACT AGACGTTGAT AATCGAGTAG TATTACCAAT AAATTCACAA ATTCGAATTT TAGTAACAGC						
Conservation							
	COX2	3,520	3,540	3,560	3,580	3,600	
C. pacifica	AGCTGACGTA ATTCATTCAT GAACTATTCC AGCTTTAGGA GTTAAGGTAG ATGGTACACC AGGACGATTA AATCAAACCT ATTTCTTAAT TAATCGACCT						3599
Translation ORF/CDS	A D V H S W T P A L G V K V D G T P G R L N Q T N F L N R P						
	COX2						
Mega1							3599
Translation ORF/CDS	A D V H S W T P A L G V K V D G T P G R L N Q T N F L N R P						
	COX2						
Mega2							3600
Translation ORF/CDS	A D V H S W T P A L G V K V D G T P G R L N Q T N F L N R P						
Consensus	AGCTGACGTA ATTCATTCAT GAACTATTCC AGCTTTAGGA GTTAAGGTAG ATGGTACACC AGGACGATTA AATCAAACCT ATTTCTTAAT TAATCGACCT						
Conservation

	ATP6	4,320	4,340	4,360	4,380	4,400	
<i>C. pacifica</i>	TAGGTCCTAT AGGAGCTAAT GGATCAACTT TCATTTTCGT TTCATTATTT TCTATAATTT TATTTAATAA TTTTATAGGA TTATTCCTT ACATTTTAC						4299
Translation ORF/CDS							
	ATP6						
<i>Mega1</i>	C.....						4299
Translation ORF/CDS							
	ATP6						
<i>Mega2</i>	...C.....C...						4300
Translation ORF/CDS							
Consensus	TAGG-CCTAT AGGAGCTAAT GGATCAACTT TCATTTTCGT TTCATTATTT TCTATAATTT TATTTAATAA TTTTATAGGA TTATT-CCTT ACATTTTAC						
Conservation	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>

	COX3	4,820	4,840	4,860	4,880	4,900	
<i>C. pacifica</i>	TGATTCCATC AATATGATGT TTCACTATTC TTTTATAGGAA ATATTATTAC TATTTTAACT GTTTATCAAT GATGACGAGA TGTTTCACGA GAAGGAACAT						4899
Translation ORF/CDS	W F H Q Y D V S L F F L G N I T I L T V Y Q W W R D V S R E G T						
	COX3						
Mega1							4899
Translation ORF/CDS	W F H Q Y D V S L F F L G N I T I L T V Y Q W W R D V S R E G T						
	COX3						
Mega2							4900
Translation ORF/CDS	W F H Q Y D V S L F F L G N I T I L T V Y Q W W R D V S R E G T						
Consensus	TGATTCCATC AATATGATGT TTCACTATTC TTTTATAGGAA ATATTATTAC TATTTTAACT GTTTATCAAT GATGACGAGA TGTTTCACGA GAAGGAACAT						
Conservation							
	COX3	4,920	4,940	4,960	4,980	5,000	
<i>C. pacifica</i>	TCCAAGGTCT TCATACTGAT GCAGTAACTA CAGGTTTACG ATGAGGAATA ATCTTATTTA TTCTTTCTGA AGTTTTATTT TTCGTCTCAT TCTTTTGAGC						4999
Translation ORF/CDS	F Q G L H T D A V T T G L R W G M I L F I L S E V L F F V S F F W A						
	COX3						
Mega1	C.....						4999
Translation ORF/CDS	F Q G L H T D A V T T G L R W G M I L F I L S E V L F F V S F F W A						
	COX3						
Mega2	C.....T.....						5000
Translation ORF/CDS	F Q G L H T D A V T T G L R W G M I L F I L S E V L F F V S F F W A						
Consensus	TCCAAGGTCT TCATACTGA GCAGTAACTA CAGGTTTACG ATGAGGAATA ATCTTATTTA TTCTTTCTGA AGTTTTATTT TTCGTCTCAT TCTTTTGAGC						
Conservation							
	COX3	5,020	5,040	5,060	5,080	5,100	
<i>C. pacifica</i>	TTTCTTCCAT AGTAGTCTTT CTCCTTCAAT TGAAGTAGGA GCTATATGAC CTCCTATAGG AATTACACCT TTTAATCCAT TCCAAATTCC TTTATTAAAT						5099
Translation ORF/CDS	F F H S S L S P S I E L G A M W P P M G I T P F N P F Q I P L L N						
	COX3						
Mega1							5099
Translation ORF/CDS	F F H S S L S P S I E L G A M W P P M G I T P F N P F Q I P L L N						
	COX3						
Mega2	G.....						5100
Translation ORF/CDS	F F H S S L S P S I E L G A M W P P M G I T P F N P F Q I P L L N						
Consensus	TTTCTTCCAT AGTAGTCTTT CTCCTTCAAT TGAAGTAGGA GCTATATGAC CTCCTATAGG AATTACACCT TTTAATCCAT TCCAAATTCC TTTATTAAAT						
Conservation							

	COX3	5,120	5,140	5,160	5,180	5,200	
<i>C. pacifica</i>	ACAGTAATTT TATTAAC TTC AGGAATTACA GTAAC TTGAG CTCATCATAG TTAAATAGAG GGAAATCATT CACAAACTAC ACAAGGATTA TTTTACAG						5199
Translation ORF/CDS	T V L L T S G T V T W A H H S L M E G N H S Q T T Q G L F F T						
	COX3						
Mega1							5199
Translation ORF/CDS	T V L L T S G T V T W A H H S L M E G N H S Q T T Q G L F F T						
	COX3						
Mega2							5200
Translation ORF/CDS	T V L L T S G T V T W A H H S L M E G N H S Q T T Q G L F F T						
Consensus	ACAGTAATTT TATTAAC TTC AGGAATTACA GTAAC TTGAG CTCATCATAG TTAAATAGAG GGAAATCATT CACAAACTAC ACAAGGATTA TTTTACAG						
Conservation							
	COX3	5,220	5,240	5,260	5,280	5,300	
<i>C. pacifica</i>	TAATCTTAGG AGTTTATTTT ACAATTTTAC AAGCTTACGA ATATATTGAA GCTCCATTTA CAATTGCAGA CTCTGTTTAT GGATCAACAT TCTTTATAGC						5299
Translation ORF/CDS	V L G V Y F T L Q A Y E Y E A P F T A D S V Y G S T F F M A						
	COX3						
Mega1	G						5299
Translation ORF/CDS	V L G V Y F T L Q A Y E Y E A P F T A D S V Y G S T F F M A						
	COX3						
Mega2							5300
Translation ORF/CDS	V L G V Y F T L Q A Y E Y E A P F T A D S V Y G S T F F M A						
Consensus	TAATCTTAGG AGTTTATTTT ACAATTTTAC AAGCTTACGA ATATATTGAA GCTCCATTTA CAATTGCAGA CTCTGTTTAT GGATCAACAT TCTTTAT GC						
Conservation							
	COX3	5,320	5,340	5,360	5,380	5,400	
<i>C. pacifica</i>	AACAGGATTC CATGGAATTC ATGTTTTAAT TGGAAC TACT TTCCTTTTAG TTTGTTTAAAT TCGACACTTA AATAATCATT TCTCTAAGAA TCATCATTTTC						5399
Translation ORF/CDS	T G F H G H V L G T T F L L V C L R H L N N H F S K N H H F						
	COX3						
Mega1							5399
Translation ORF/CDS	T G F H G H V L G T T F L L V C L R H L N N H F S K N H H F						
	COX3						
Mega2							5400
Translation ORF/CDS	T G F H G H V L G T T F L L V C L R H L N N H F S K N H H F						
Consensus	AACAGGATTC CATGGAATTC ATGTTTTAAT TGGAAC TACT TTCCTTTTAG TTTGTTTAAAT TCGACACTTA AATAATCATT TCTCTAAGAA TCATCATTTTC						
Conservation

		6,320	6,340	6,360	6,380	ND5	
<i>C. pacifica</i>	TCTTCAGTGT CATACTCTAA ATATAAGCTA TTTGAATAAA AACAAATTAT ATACATATAT AAAATTATAA TTCAAAAAAT AAAACTTAAT AAATAAATCT					6397	
Translation ORF/CDS			F C M Y M Y L M W F F S L L Y K				
		6,320	6,340	6,360	6,380	ND5	
Mega1	G.....					6399	
Translation ORF/CDS			F C M Y M Y L M W F F S L L Y K				
		6,320	6,340	6,360	6,380	ND5	
Mega2						6398	
Translation ORF/CDS			F C M Y M Y L M W F F S L L Y K				
Consensus	TCTTCAGTGT CATACTCTAA ATATAAGCTA TTTGAATAAA AACAAATTAT ATACATATAT AAAATTATAA TTCAAAAAAT AAAACTTAAT AAATAAATCT						
Conservation							
		6,420	6,440	6,460	6,480	ND5	
<i>C. pacifica</i>	TTAAATTATT ATTTTGTAAT ATTTGATTTA ATTGAGAATT CTTTACTAAA TTATAATAAA GTTGTGGCCC CCCAAAATAT TCTGATCATC CTTGATCAAA					6497	
Translation ORF/CDS	L N N N Q L M Q N L Q S N K V L N Y Y L Q Q G G F Y E S W G Q D F						
		6,420	6,440	6,460	6,480	ND5	
Mega1						6499	
Translation ORF/CDS	L N N N Q L M Q N L Q S N K V L N Y Y L Q Q G G F Y E S W G Q D F						
		6,420	6,440	6,460	6,480	ND5	
Mega2	G.....					6498	
Translation ORF/CDS	L N N N Q L M Q N L Q S N K V L N Y Y L Q Q G G F Y E S W G Q D F						
Consensus	TTAAATTATT ATTTTGTAAT ATTTGATTTA ATTGAGAATT CTTTACTAAA TTATAATAAA GTTGTGGCCC CCCAAAATAT TCTGATCATC CTTGATC-AA						
Conservation							
		6,520	6,540	6,560	6,580	ND5	
<i>C. pacifica</i>	TGACTTAACT ACCATCTTTC CTACAATTAA TGAATAATTA ATAATTCCAT AAGTAGAAAT ATATGGTATA AATCACATTG AACCTAAAAA ATAAGAAGAA					6597	
Translation ORF/CDS	S K V V M K G V L S Y N I I G Y T S I Y P M F W M S G L F Y S S						
		6,520	6,540	6,560	6,580	ND5	
Mega1						6599	
Translation ORF/CDS	S K V V M K G V L S Y N I I G Y T S I Y P M F W M S G L F Y S S						
		6,520	6,540	6,560	6,580	ND5	
Mega2						6598	
Translation ORF/CDS	S K V V M K G V L S Y N I I G Y T S I Y P M F W M S G L F Y S S						
Consensus	TGACTTAACT ACCATCTTTC CTACAATTAA TGAATAATTA ATAATTCCAT AAGTAGAAAT ATATGGTATA AATCACATTG AACCTAAAAA ATAAGAAGAA						
Conservation							

		5,620	5,640	5,660	5,680	ND5
C. pacifica	TTATAATTAT TTAAAGCCTT ATTATAAAAA AACAGAGACA CATTAGAAAT TATGTAACCT CTTACCCCTC CTACAATACA AACAAATAAT GTTAACAAC	6697				
Translation ORF/CDS	N Y N N L A K N Y F F L S V N S I M Y G S V G G V I C V F L T L L K					
		ND5				
Mega1		6699				
Translation ORF/CDS	N Y N N L A K N Y F F L S V N S I M Y G S V G G V I C V F L T L L K					
		ND5				
Mega2		6698				
Translation ORF/CDS	N Y N N L A K N Y F F L S V N S I M Y G S L G G V I C V F L T L L K					
Consensus	TTATAATTAT TTAAAGCCTT ATTATAAAAA AACAGAGACA CATTAGAAAT TATGTAACCT CTTA--CCTC CTACAATACA AACAAATAAT GTTAACAAC					
Conservation						
		5,720	5,740	5,760	5,780	ND5
C. pacifica	TTAAATAAGA AGGAAGAACT ACTACTACTG GAGTAGGAAA AATTAATCAA CTTAGTATTC TACCTCCAAA AATTCTTAAA ATTAATAATC CTATTATACT	6797				
Translation ORF/CDS	L Y S P L V V V V P T P F I L W S L M S G G F I S L I L L G M M S					
		ND5				
Mega1		6799				
Translation ORF/CDS	L Y S P L V V V V P T P F I L W S L M S G G F I S L I L L G M M S					
		ND5				
Mega2		6798				
Translation ORF/CDS	L Y S P L V V V V P T P F I L W S L M S G G F I S L I L L G M M S					
Consensus	TTAAATAAGA AGGAAGAACT ACTACTACTG GAGTAGGAAA AATTAATCAA CTTAGTATTC TACCTCCAAA AATTCTTAAA ATTAATAATC CTATTATACT					
Conservation						
		5,820	5,840	5,860	5,880	ND5
C. pacifica	CTTTAACATT ACTCAACCTT CATCATTTAA TATATTTAAA GAAGAAAAAT TTGAATCTCC TGTATTGTA TAATAAACTA ATCGAAAAGA ATAACAAACA	6897				
Translation ORF/CDS	K L M V W G E D N L M N L S S F N S D G T M T Y Y V L R F S Y C V					
		ND5				
Mega1		6899				
Translation ORF/CDS	K L M V W G E D N L M N L S F F N S D G T M T Y Y V L R F S Y C V					
		ND5				
Mega2		6898				
Translation ORF/CDS	K L M V W G E D N L M N L S S F N S D G T M T Y Y V L R F S Y C V					
Consensus	CTTTAACATT ACTCAACCTT CATCATTTAA TATATTTAAA GAA-AAAAAT TTGA-TCTCC TGTATTGTA TAATAAACTA ATCGAAAAGA ATAACAAACA					
Conservation						

		6,520	6,540	6,560	6,580	6,600	6,620	6,640	6,660	6,680	6,700	6,720	6,740	6,760	6,780	6,800	6,820	6,840	6,860	6,880	6,900	6,920	6,940	6,960	6,980	7,000	7,020	7,040	7,060	7,080	7,100	7,120	7,140	7,160	7,180	7,200	7,220	7,240	7,260	7,280	7,300	7,320	7,340	7,360	7,380	7,400	7,420	7,440	7,460	7,480	7,500	7,520	7,540	7,560	7,580	7,600	7,620	7,640	7,660	7,680	7,700	7,720	7,740	7,760	7,780	7,800	7,820	7,840	7,860	7,880	7,900	7,920	7,940	7,960	7,980	8,000	8,020	8,040	8,060	8,080	8,100	8,120	8,140	8,160	8,180	8,200	8,220	8,240	8,260	8,280	8,300	8,320	8,340	8,360	8,380	8,400	8,420	8,440	8,460	8,480	8,500	8,520	8,540	8,560	8,580	8,600	8,620	8,640	8,660	8,680	8,700	8,720	8,740	8,760	8,780	8,800	8,820	8,840	8,860	8,880	8,900	8,920	8,940	8,960	8,980	9,000	9,020	9,040	9,060	9,080	9,100	9,120	9,140	9,160	9,180	9,200	9,220	9,240	9,260	9,280	9,300	9,320	9,340	9,360	9,380	9,400	9,420	9,440	9,460	9,480	9,500	9,520	9,540	9,560	9,580	9,600	9,620	9,640	9,660	9,680	9,700	9,720	9,740	9,760	9,780	9,800	9,820	9,840	9,860	9,880	9,900	9,920	9,940	9,960	9,980	10,000	10,020	10,040	10,060	10,080	10,100	10,120	10,140	10,160	10,180	10,200	10,220	10,240	10,260	10,280	10,300	10,320	10,340	10,360	10,380	10,400	10,420	10,440	10,460	10,480	10,500	10,520	10,540	10,560	10,580	10,600	10,620	10,640	10,660	10,680	10,700	10,720	10,740	10,760	10,780	10,800	10,820	10,840	10,860	10,880	10,900	10,920	10,940	10,960	10,980	11,000	11,020	11,040	11,060	11,080	11,100	11,120	11,140	11,160	11,180	11,200	11,220	11,240	11,260	11,280	11,300	11,320	11,340	11,360	11,380	11,400	11,420	11,440	11,460	11,480	11,500	11,520	11,540	11,560	11,580	11,600	11,620	11,640	11,660	11,680	11,700	11,720	11,740	11,760	11,780	11,800	11,820	11,840	11,860	11,880	11,900	11,920	11,940	11,960	11,980	12,000	12,020	12,040	12,060	12,080	12,100	12,120	12,140	12,160	12,180	12,200	12,220	12,240	12,260	12,280	12,300	12,320	12,340	12,360	12,380	12,400	12,420	12,440	12,460	12,480	12,500	12,520	12,540	12,560	12,580	12,600	12,620	12,640	12,660	12,680	12,700	12,720	12,740	12,760	12,780	12,800	12,820	12,840	12,860	12,880	12,900	12,920	12,940	12,960	12,980	13,000	13,020	13,040	13,060	13,080	13,100	13,120	13,140	13,160	13,180	13,200	13,220	13,240	13,260	13,280	13,300	13,320	13,340	13,360	13,380	13,400	13,420	13,440	13,460	13,480	13,500	13,520	13,540	13,560	13,580	13,600	13,620	13,640	13,660	13,680	13,700	13,720	13,740	13,760	13,780	13,800	13,820	13,840	13,860	13,880	13,900	13,920	13,940	13,960	13,980	14,000	14,020	14,040	14,060	14,080	14,100	14,120	14,140	14,160	14,180	14,200	14,220	14,240	14,260	14,280	14,300	14,320	14,340	14,360	14,380	14,400	14,420	14,440	14,460	14,480	14,500	14,520	14,540	14,560	14,580	14,600	14,620	14,640	14,660	14,680	14,700	14,720	14,740	14,760	14,780	14,800	14,820	14,840	14,860	14,880	14,900	14,920	14,940	14,960	14,980	15,000	15,020	15,040	15,060	15,080	15,100	15,120	15,140	15,160	15,180	15,200	15,220	15,240	15,260	15,280	15,300	15,320	15,340	15,360	15,380	15,400	15,420	15,440	15,460	15,480	15,500	15,520	15,540	15,560	15,580	15,600	15,620	15,640	15,660	15,680	15,700	15,720	15,740	15,760	15,780	15,800	15,820	15,840	15,860	15,880	15,900	15,920	15,940	15,960	15,980	16,000	16,020	16,040	16,060	16,080	16,100	16,120	16,140	16,160	16,180	16,200	16,220	16,240	16,260	16,280	16,300	16,320	16,340	16,360	16,380	16,400	16,420	16,440	16,460	16,480	16,500	16,520	16,540	16,560	16,580	16,600	16,620	16,640	16,660	16,680	16,700	16,720	16,740	16,760	16,780	16,800	16,820	16,840	16,860	16,880	16,900	16,920	16,940	16,960	16,980	17,000	17,020	17,040	17,060	17,080	17,100	17,120	17,140	17,160	17,180	17,200	17,220	17,240	17,260	17,280	17,300	17,320	17,340	17,360	17,380	17,400	17,420	17,440	17,460	17,480	17,500	17,520	17,540	17,560	17,580	17,600	17,620	17,640	17,660	17,680	17,700	17,720	17,740	17,760	17,780	17,800	17,820	17,840	17,860	17,880	17,900	17,920	17,940	17,960	17,980	18,000	18,020	18,040	18,060	18,080	18,100	18,120	18,140	18,160	18,180	18,200	18,220	18,240	18,260	18,280	18,300	18,320	18,340	18,360	18,380	18,400	18,420	18,440	18,460	18,480	18,500	18,520	18,540	18,560	18,580	18,600	18,620	18,640	18,660	18,680	18,700	18,720	18,740	18,760	18,780	18,800	18,820	18,840	18,860	18,880	18,900	18,920	18,940	18,960	18,980	19,000	19,020	19,040	19,060	19,080	19,100	19,120	19,140	19,160	19,180	19,200	19,220	19,240	19,260	19,280	19,300	19,320	19,340	19,360	19,380	19,400	19,420	19,440	19,460	19,480	19,500	19,520	19,540	19,560	19,580	19,600	19,620	19,640	19,660	19,680	19,700	19,720	19,740	19,760	19,780	19,800	19,820	19,840	19,860	19,880	19,900	19,920	19,940	19,960	19,980	20,000	20,020	20,040	20,060	20,080	20,100	20,120	20,140	20,160	20,180	20,200	20,220	20,240	20,260	20,280	20,300	20,320	20,340	20,360	20,380	20,400	20,420	20,440	20,460	20,480	20,500	20,520	20,540	20,560	20,580	20,600	20,620	20,640	20,660	20,680	20,700	20,720	20,740	20,760	20,780	20,800	20,820	20,840	20,860	20,880	20,900	20,920	20,940	20,960	20,980	21,000	21,020	21,040	21,060	21,080	21,100	21,120	21,140	21,160	21,180	21,200	21,220	21,240	21,260	21,280	21,300	21,320	21,340	21,360	21,380	21,400	21,420	21,440	21,460	21,480	21,500	21,520	21,540	21,560	21,580	21,600	21,620	21,640	21,660	21,680	21,700	21,720	21,740	21,760	21,780	21,800	21,820	21,840	21,860	21,880	21,900	21,920	21,940	21,960	21,980	22,000	22,020	22,040	22,060	22,080	22,100	22,120	22,140	22,160	22,180	22,200	22,220	22,240	22,260	22,280	22,300	22,320	22,340	22,360	22,380	22,400	22,420	22,440	22,460	22,480	22,500	22,520	22,540	22,560	22,580	22,600	22,620	22,640	22,660	22,680	22,700	22,720	22,740	22,760	22,780	22,800	22,820	22,840	22,860	22,880	22,900	22,920	22,940	22,960	22,980	23,000	23,020	23,040	23,060	23,080	23,100	23,120	23,140	23,160	23,180	23,200	23,220	23,240	23,260	23,280	23,300	23,320	23,340	23,360	23,380	23,400	23,420	23,440	23,460	23,480	23,500	23,520	23,540	23,560	23,580	23,600	23,620	23,640	23,660	23,680	23,700	23,720	23,740	23,760	23,780	23,800	23,820	23,840	23,860	23,880	23,900	23,920	23,940	23,960	23,980	24,000	24,020	24,040	24,060	24,080	24,100	24,120	24,140	24,160	24,180	24,200	24,220	24,240	24,260	24,280	24,300	24,320	24,340	24,360	24,380	24,400	24,420	24,440	24,460	24,480	24,500	24,520	24,540	24,560	24,580	24,600	24,620	24,640	24,660	24,680	24,700	24,720	24,740	24,760	24,780	24,800	24,820	24,840	24,860	24,880	24,900	24,920	24,940	24,960	24,980	25,000	25,020	25,040	25,060	25,080	25,100	25,120	25,140	25,160	25,180	25,200	25,220	25,240	25,260	25,280	25,300	25,320	25,340	25,360	25,380	25,400	25,420	25,440	25,460	25,480	25,500	25,520	25,540	25,560	25,580	25,600	25,620	25,640	25,660	25,680	25,700	25,720	25,740	25,760	25,780	25,800	25,820	25,840	25,860	25,880	25,900	25,920	25,940	25,960	25,980	26,000	26,020	26,040	26,060	26,080	26,100	26,120	26,140	26,160	26,180	26,200	26,220	26,240	26,260	26,280	26,300	26,320	26,340	26,360	26,380	26,400	26,420	26,440	26,460	26,480	26,500	26,520	26,540	26,560	26,580	26,600	26,620	26,640	26,660	26,680	26,700	26,720	26,740	26,760	26,780	26,800	26,820	26,840	26,860	26,880	26,900	26,920	26,940	26,960	26,980	27,000	27,020	27,040	27,060	27,080	27,100	27,120	27,140	27,160	27,180	27,200	27,220	27,240	27,260	27,280	27,300	27,320	27,340	27,360	27,380	27,400	27,420	27,440	27,460	27,480	27,500	27,520	27,540	27,560	27,580	27,600	27,620	27,640	27,660	27,680	27,700	27,720	27,740	27,760	27,780	27,800	27,820	27,840	27,860	27,880	27,900	27,920	27,940	27,960	27,980	28,000	28,020	28,040	28,060	28,080	28,100	28,120	28,140	28,160	28,180	28,200	28,220	28,240	28,260	28,280	28,300	28,320	28,340	28,360	28,380	28,400	28,420	28,440	28,460	28,480	28,500	28,520	28,540	28,560	28,580	28,600	28,620	28,640	28,660	28,680	28,700	28,720	28,740	28,760	28,780	28,800	28,820	28,840	28,860	28,880	28,900	28,920	28,940	28,960	28,980	29,000	29,020	29,040	29,060	29,080	29,100	29,120	29,140	29,160	29,180	29,200	29,220	29,240	29,260
--	--	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

		7,320	7,340	7,360	7,380	ND5
C. pacifica	ATTGACAAAA TTCTCATTAT TAATCCTAAT TGACTTAAAG TAGACAAAGC AATAATCTTC TTAAATCAA ATTCATAATT AGCTCCTAAT CCAGCTATAA					7297
Translation ORF/CDS	M S L I S M M L G L Q S L T S L A I I K K L D F E Y N A G L G A M F					ND5
Mega1						7299
Translation ORF/CDS	M S L I S M M L G L Q S L T S L A I I K K L D F E Y N A G L G A M F					ND5
Mega2						7298
Translation ORF/CDS	M S L I S M M L G L Q S L T S L A I I K K L D F E Y N A G L G A M F					
Consensus	ATTGACAAAA TTCT-ATTAT TAATCCTAAT TG-CTTAAAG TAGACAAAGC AATAATCTTC TTAAATCAA ATTCATAATT AGCTCCTAAT CCAGCTATAA					
Conservation						
		7,320	7,340	7,360	7,380	ND5
C. pacifica	ATATTGTTAA TCCAGATAAT AATAATAATA AATTACCCAT TCATGAATCA CTTAATAAAA TATTAACCG AATTAATAAA TAAACCCCTG CAGTTACCAA					7397
Translation ORF/CDS	M T L G S L L L L L N G M W S D S L L I N F R I L L Y V G A T V L					ND5
Mega1						7399
Translation ORF/CDS	M T L G S L L L L L N G M W S D S L L I N F R I L L Y V G A T V L					ND5
Mega2						7398
Translation ORF/CDS	M T L G S L L L L L N G M W S D S L L I N F R I L L Y V G A T V L					
Consensus	ATATTGTTAA TCCAGATAAT AATAATAATA AATTACCCAT TCATGAATCA CTTAATAAAA TATTAACCG AATTAATAAA TAAACCCCTG CAGTTACCAA					
Conservation						
		7,420	7,440	7,460	7,480	ND5
C. pacifica	AGTAGAAGAA TGAAGTAAAG CAGAAACAGG AGTTGGAGCT GCTATAGCAG CTGGTAATCA TGAAGAAAAA GGAATCTGTG CTCTCTTTGT TATTGCAGCC					7497
Translation ORF/CDS	T S S H V L A S V P T P A A M A A P L W S S F P I Q A S K T M A A					ND5
Mega1						7499
Translation ORF/CDS	T S S H V L A S V P T P A A M A A P L W S S F P I Q A S K T M A A					ND5
Mega2						7498
Translation ORF/CDS	T S S H V L A S V P T P A A M A A P L W S S F P I Q A S K T M A A					
Consensus	AGTAGAAGAA TGAAGTAAAG CAGAAACAGG AGTTGGAGCT GCTATAGCAG CTGGTAATCA TGAAGAAAAA GGAAT-TGTG CTCTCTTTGT TATTGCAGCC					
Conservation						

		7,520		7,540		7,560		7,580		ND5																										
C. pacifica	AACATTACTA	AAC	TACCAAC	AAGTATTATT	TCAAAATCAT	TCTTTATAAC	TTCTAAATAA	AAAACATAAT	TTCAACTTCC	ATAATTTAAT	ATTCATGCAA	7597																								
Translation ORF/CDS	L	M	V	L	S	G	V	L	M	M	E	F	D	N	K	M	V	E	L	Y	F	V	Y	N	W	S	G	Y	N	L	M	W	A	I	ND5	
Mega1																																			7599	
Translation ORF/CDS	L	M	V	L	S	G	V	L	M	M	E	F	D	N	K	M	V	E	L	Y	F	V	Y	N	W	S	G	Y	N	L	M	W	A	I	ND5	
Mega2																																			7598	
Translation ORF/CDS	L	M	V	L	S	G	V	L	M	M	E	F	D	N	K	M	V	E	L	Y	F	V	Y	N	W	S	G	Y	N	L	M	W	A	I	ND5	
Consensus	AACATTACTA	AAC	TACCAAC	AAGTATTATT	TCAAAATCAT	TCTTTATAAC	TTCTAAATAA	AAAACATAAT	TTCAACTTCC	ATAATTTAAT	ATTCATGCAA																									
Conservation																																				
		7,620		7,640		7,660		7,680		ND5																										
C. pacifica	TAGCCAATAA	TAAAGCAACA	TCACCAATCC	GATTTGATAA	AGCAGTTAAT	ATACCAGCAT	TATAAGACTT	AACATTTTGA	AAATAAATAA	CCAAACAATA	7697																									
Translation ORF/CDS	A	L	L	L	A	V	D	G	I	R	N	S	L	A	T	L	M	G	A	N	Y	S	K	V	N	Q	F	Y	I	V	L	C	Y	ND5		
Mega1																																			7699	
Translation ORF/CDS	A	L	L	L	A	V	D	G	I	R	N	S	L	A	T	L	M	G	A	N	Y	S	K	V	N	Q	F	Y	I	V	L	C	Y	ND5		
Mega2																																			7698	
Translation ORF/CDS	A	L	L	L	A	V	D	G	I	R	N	S	L	A	T	L	M	G	A	N	Y	S	K	V	N	Q	F	Y	I	V	L	C	Y	ND5		
Consensus	TAGCCAATAA	TAAAGCAACA	TCACCAATCC	GATTTGATAA	AGCAGTTAAT	ATACCAGCAT	TATAAGACTT	AACATTTTGA	AAATAAATAA	CCAAACAATA																										
Conservation																																				
		7,720		7,740		7,760		7,780		ND5																										
C. pacifica	AGAAACAAGT	CCTAACCCAT	CTCATCCTAA	TAAAATACTA	ATTAAATTAG	GACTAATAAT	TAATAACATC	ATTGATGTTA	CAACATTAA	AACTAATATA	7797																									
Translation ORF/CDS	S	V	L	G	L	G	D	W	G	L	L	I	S	I	L	N	P	S	I	I	L	L	M	M	S	T	V	F	M	L	V	L	M	ND5		
Mega1																																			7799	
Translation ORF/CDS	S	V	L	G	L	G	D	W	G	L	L	I	S	I	L	N	P	S	I	I	L	L	M	M	S	T	V	F	M	L	V	L	M	ND5		
Mega2																																			7798	
Translation ORF/CDS	S	V	L	G	L	G	D	W	G	L	L	I	S	I	L	N	P	S	I	I	L	L	M	M	S	T	V	F	M	L	V	L	M	ND5		
Consensus	AGAAACAAGT	CCTAACCCAT	CTCATCCTAA	TAAAATACTA	AT	AAATTAG	GACTAATAAT	TAATAACAT	ATTGATGTTA	CAACATTAA	AACTAATATA																									
Conservation

		8,420	8,440	8,460	8,480	ND4
C. pacifica	GGTGGAGCAG CCATATTAGC AGAACTTAAT AAAAATCATC ATAATGCTAT TGAAGGTATA AAATTTAATA ATCCCTTATT AATTAATAAA CTTCGACTTC					8497
Translation ORF/CDS	P P A A M N A S S L L F W W L A M S P M F N L L G K N I L L S R S G					
						ND4
Mega1						8499
Translation ORF/CDS						
						ND4
Mega2	C.....					8498
Translation ORF/CDS	P P A A M N A S S L L F W W L A M S P M F N L L G K N I L L S R S G					
Consensus	GGTGGAGCAG CCATATTAGC AGAACTTAAT AAAAATCATC ATAATGCTAT TGAAGGTATA AAATTTAATA A-CCCTTATT AATTAATAAA CTTCGACTTC					
Conservation						
		8,520	8,540	8,560	8,580	ND4
C. pacifica	CCAATCGTTC ATAAGAAATA TTAGCTAAAC AAAAATATCC TGAAGAACAT AATCCATGAG CAATTATTAA AGTATAAGAT CCACAAATTC CTATATATGT					8597
Translation ORF/CDS	L R E Y S I N A L C F L G S S C L G H A I M L T Y S G C I G M Y T					
						ND4
Mega1						8599
Translation ORF/CDS						
						ND4
Mega2	C.....C.....					8598
Translation ORF/CDS	L R E Y S I N A L C F L G S S C L G H A I M L T Y S G C I G M Y T					
Consensus	CCAATCGTTC ATAAGAAATA TTAGCTAAAC AAAA-AATCC TGAAGAACAT AATCCATGAG CAATTATTAA AGTATAAGAT CCACAAATTC C-ATATATGT					
Conservation						
		8,620	8,640	8,660	8,680	ND4
C. pacifica	TAAAGTTATT AAACCAGCTA ATACAATTCC TATATGAGCA ACAGAAGAAT AAGCAATCAA TGCCTTTAAA TCTGTTTGTC GTAAACAAAT TAATCTTACT					8697
Translation ORF/CDS	L T M L G A L V I G M H A V S S Y A I L A K L D T Q R L C I L S V					
						ND4
Mega1						8699
Translation ORF/CDS						
						ND4
Mega2						8698
Translation ORF/CDS	L T M L G A L V I G M H A V S S Y A I L A K L D T Q R L C I L S V					
Consensus	TAAAGTTATT AAACCAGCTA ATACAATTCC TATATGAGCA ACAGAAGAAT AAGCAATCAA TGCCTTTAAA TCTGTTTGTC GTAAACAAAT TAATCTTACT					
Conservation						

		9,120	9,140	9,160	9,180	ND4
C. pacifica	AAAATTCCAA TTAATATTGG TAAAGAACT AATAAAGTAT AAAATAATAA ATATACTCCT GCTTGAATC GTTCAGGTTG ATACCCTCAT CCTAAAATCA					9097
Translation ORF/CDS	F I G I L M P L S V L L T Y F L L Y V G A Q L R E P Q Y G W G L I L					
						ND4
Mega1						9099
Translation ORF/CDS						
						ND4
Mega2	A.....					9098
Translation ORF/CDS	F I G I L M P L S V L L T Y F L L Y V G A Q L R E P Q Y G W G L I L					
Consensus	AAAAT-CCAA TTAATATTGG TAAAGAACT AATAAAGTAT AAAATAATAA ATATACTCCT GCTTGAATC GTTCAGGTTG ATACCCTCAT CCTAAAATCA					
Conservation						
		9,120	9,140	9,160	9,180	ND4
C. pacifica	AAAACAATGT AGGAATTAAA CTACTTTCAA AAAATAAATA AAATATAAAT AAACCTATAC TTCTAAAAGT TAACACTAAT AAAATCAATA ATAAAATAAT					9197
Translation ORF/CDS	F L T P I L S S E F F L Y F M F L S M S S F T L V L L I L L L I I					
						ND4
Mega1						9199
Translation ORF/CDS						
						ND4
Mega2	T.....					9198
Translation ORF/CDS	F L T P I L S S E F F L Y F M F L S M S S F T L V L L I L L L I I					
Consensus	AAAACAATGT AGGAATTAAA CTACTTTCAA AAAATAAATA AAATATAAAT AAACCTATAC TTCTAAAAGT TAA-ACTAAT AAAATCAATA ATAAAATAAT					
Conservation						
		9,220	9,240	9,260	9,280	ND4
C. pacifica	ATTTAATAAA AACAAATTAG TATAATTTCT ATACTTATAA ATTGATTCAC TAGCCATCAA CATTAAAGAA ACAATTCATA ATCTTAATAA AATTAATCCA					9297
Translation ORF/CDS	N L L F L N T Y N S Y K Y I S E S A M L M L S V I W L S L L I L G					
						ND4
Mega1						9299
Translation ORF/CDS						
						ND4
Mega2						9298
Translation ORF/CDS	N L L F L N T Y N S Y K Y I S E S A M L M L S V I W L S L L I L G					
Consensus	ATTTAATAAA AACAAATTAG TATAATTTCT ATACTTATAA ATTGATTCAC TAGCCATCAA CATTAAAGAA ACAATTCATA ATCTTAATAA AATTAATCCA					
Conservation

	ND6	10,220	10,240	10,260	10,280	10,300	
<i>C. pacifica</i>	CTTCATAGAT AAAAATTTAC TTCTTCAATA CCAAAATTTA GAAGTAAAGT CAATTTATGA CATAAATTCA TATTTAATAG AAAATTCATT ATCCCTTAAT						10297
Translation ORF/CDS							
	ND6						
Mega1							10299
Translation ORF/CDS							
	ND6						
Mega2						C	10298
Translation ORF/CDS	F M D K N L L L Q Y Q N L E V K S I Y D M N S Y L M E N S L S L N						
Consensus	CTTCATAGAT AAAAATTTAC TTCTTCAATA CCAAAATTTA GAAGTAAAGT CAATTTATGA CATAAATTCA TAT-TAATAG AAAATTCATT ATCCCTTAAT						
Conservation							
	ND6	10,320	10,340	10,360	10,380	10,400	
<i>C. pacifica</i>	AAATTATATA ACTACCCAAC AAATTTACTA ACAATTTTAT TAATAAATTA CTTATTAATT ACATTAATTG CAGTAGTAAA AATTACAAAG TTATTTAAGG						10397
Translation ORF/CDS							
	ND6						
Mega1							10399
Translation ORF/CDS							
	ND6						
Mega2							10398
Translation ORF/CDS	K L Y N Y P T N L L T I L L M N Y L L I T L I A V V K I T K L F K						
Consensus	AAATTATATA ACTACCCAAC AAATTTACTA ACAATTTTAT TAATAAATTA CTTATTAATT ACATTAATTG CAGTAGTAAA AATTACAAAG TTATTTAAGG						
Conservation							
	ND6	10,420	10,440	10,460	10,480	10,500	
<i>C. pacifica</i>	GACCTCTTCG ACCAATATTT AACTAATGAA CAAACCTTTA CGAGTTAAAC ACCCAATTTT CAGAATTGCT AATAGTGCTT TAGTAGATTT ACCTGCTCCT						10497
Translation ORF/CDS	M N K P L R V K H P I F S I A N S A L V D L P A P						
	ND6						
Mega1							10499
Translation ORF/CDS	M N K P L R V K H P I F S I A N S A L V D L P A P						
	ND6						
Mega2							10498
Translation ORF/CDS	G P L R P M F N *						
Consensus	GACCTCTTCG ACCAATATTT AACTAATGAA CAAACCTTTA CGAGTTAAAC ACCCAATTTT CAGAATTGCT AATAGTGCTT TAGTAGATTT ACCTGCTCCT						
Conservation							

	CYTb	10,520	10,540	10,560	10,580	10,600	
C. pacifica	AGAAATATTT CAGCCTGATG AAATTTTGGG TCATTATTAT TCTTATGTTT AATGATTCAA ATTTTAACTG GATTATTTTT AGCTATACAT TACACAGCAG						10597
Translation ORF/CDS	S N S A W W N F G S L L F L C L M Q L T G L F L A M H Y T A						
	CYTb						
Mega1							10599
Translation ORF/CDS	S N S A W W N F G S L L F L C L M Q L T G L F L A M H Y T A						
	CYTb						
Mega2							10598
Translation ORF/CDS							
Consensus	AGAAATATTT CAGCCTGATG AAATTTTGGG TCATTATTAT TCTTATGTTT AATGATTCAA ATTTTAACTG GATTATTTTT AGCTATACAT TACACAGCAG						
Conservation							
	CYTb	10,620	10,640	10,660	10,680	10,700	
C. pacifica	ATATTAATTT AGCTTTTAAT AGAGTAAATC ATATTTGTCT AGATGTAAAC TATGGATGAT TATTACGAAC TATACATGCT AATGGTGAT CATTCTTCTT						10697
Translation ORF/CDS	D N L A F N S V N H C R D V N Y G W L L R T M H A N G A S F F F						
	CYTb						
Mega1	 C						10699
Translation ORF/CDS	D N L A F N S V N H C R D V N Y G W L L R T M H A N G A S F F F						
	CYTb						
Mega2	 C						10698
Translation ORF/CDS							
Consensus	ATATTAATTT AGCTTTTAAT AGAGTAAATC ATATTTG-CG AGATGTAAAC TATGGATGAT TATTACGAAC TATACATGCT AATGGTGAT CATTCTTCTT						
Conservation							
	CYTb	10,720	10,740	10,760	10,780	10,800	
C. pacifica	TATTTGTATT TATCTACATG TAGGACGAGG AATTTACTAT GGTTCTACT TATTTACTCC AACTTGATTA GTAGGAGTTA TTATTTTATT CTTAGTAATA						10797
Translation ORF/CDS	C Y L H V G R G Y Y G S Y L Y T P T W L V G V L F L V M						
	CYTb						
Mega1	 A						10799
Translation ORF/CDS	C Y L H V G R G Y Y G S Y L Y T P T W L V G V L F L V M						
	CYTb						
Mega2							10798
Translation ORF/CDS							
Consensus	TATTTGTATT TATCTACATG TAGGACGAGG AATTTACTAT GGTTCTACT TAT-TACTCC AACTTGATTA GTAGGAGTTA TTATTTTATT CTTAGTAATA						
Conservation							

	CYTb	10,920	10,940	10,960	10,980	10,995	
<i>C. pacifica</i>	GGAACAGCTT TTATAGGATA CGTTCTTCCT TGAGGACAAA TATCTTTTTG AGGAGCTACT GTTATTACTA ATTTATTATC AGCTATTCCT TATTTAGGAA						10897
Translation ORF/CDS	G T A F M G Y V L P W G Q M S F W G A T V T N L L S A P Y L G						
	CYTb						
Mega1							10899
Translation ORF/CDS	G T A F M G Y V L P W G Q M S F W G A T V T N L L S A P Y L G						
	CYTb						
Mega2	 T						10898
Translation ORF/CDS							
Consensus	GGAACAGCTT TTATAGGATA -GTTCTTCCT TGAGGACAAA TATCTTTTTG AGGAGCTACT GTTATTACTA ATTTATTATC AGCTATTCCT TATTTAGGAA						
Conservation							
	CYTb	10,920	10,940	10,960	10,980	11,000	
<i>C. pacifica</i>	TTGACCTAGT ACAATGGGTA TGAGGAGGGT TTGCTGTTGA TAACGCTACA TTAACGCGAT TCTTCACATT CCACTTCATT TTACCATTTA TTGTTCTAGC						10997
Translation ORF/CDS	I D L V Q W V W G G F A V D N A T L T R F F T F H F I L P F V L A						
	CYTb						
Mega1	 A A C						10999
Translation ORF/CDS	I D L V Q W V W G G F A V D N A T L T R F F T F H F I L P F V L A						
	CYTb						
Mega2	 T A						10998
Translation ORF/CDS							
Consensus	TTGA-CTAGT ACAATG-GTA TGAGGAGG-T TTGCTGTTGA TAACGCTACA TTAAC-CGAT TCTTCACATT CCACTTCATT TTACCATTTA TTGTTCTAGC						
Conservation							
	CYTb	11,020	11,040	11,060	11,080	11,100	
<i>C. pacifica</i>	TGCAACTCTT ATTCATATTT TATTCCTTCA CGAAACAGGA TCAAATAATC CAATAGGATT AAATTCTAAT ATTGATAAAA TTCCATTCCA TCCTTATTTT						11097
Translation ORF/CDS	A T L I H L F L H E T G S N N P M G L N S N I D K I P F H P Y F						
	CYTb						
Mega1							11099
Translation ORF/CDS	A T L I H L F L H E T G S N N P M G L N S N I D K I P F H P Y F						
	CYTb						
Mega2	 C						11098
Translation ORF/CDS							
Consensus	TGCAACTCTT ATTCATATTT TATTCCTTCA CGAAACAGGA TCAAATAATC CAATAGGATT AAATTCTAAT ATTGATAAAA TTCCATTCCA TCCTTA-TTC						
Conservation							

	CYTb	11,120	11,140	11,160	11,180	11,200	
C. pacifica	ACTTACAAGG ATATTGTAGG ATTTATTGTT ATAACAATAA TTCTAATTTT ATTAGTATTA ATTAATCCTT ACTTACTTGG TGACCCTGAT AATTTTATCC						11197
Translation ORF/CDS	T Y K D V G F V M T M L L L V L N P Y L L G D P D N F						
	CYTb						
Mega1							11199
Translation ORF/CDS	T Y K D V G F V M T M L L L V L N P Y L L G D P D N F						
	CYTb						
Mega2							11198
Translation ORF/CDS							
Consensus	ACTTACAAGG ATATTGTAGG ATTTATTGTT ATAACAATAA TTCTAATTTT ATTAGTATTA ATTAATCCTT ACTTACTTGG TGACCCTGAT AATTTTATCC						
Conservation							
	CYTb	11,220	11,240	11,260	11,280	11,300	
C. pacifica	CAGCCAATCC ATTAGTTACT CCAGTTCATA TTCAACCTGA ATGATACTTC TTATTTGCTT ACGCAATTTT ACGATCTATT CCTAATAAGT TAGGAGGAGT						11297
Translation ORF/CDS	P A N P L V T P V H Q P E W Y F L F A Y A L R S P N K L G G V						
	CYTb						
Mega1	 T						11299
Translation ORF/CDS	P A N P L V T P V H Q P E W Y F L F A Y A L R S P N K L G G V						
	CYTb						
Mega2	. G T						11298
Translation ORF/CDS							
Consensus	C-GCCAATCC -TTAGTTACT CCAGTTCATA TTCAACCTGA ATGATACTTC TTATTTGCTT ACGCAATTTT ACGATCTATT CCTAATAA-T TAGGAGGAGT						
Conservation							
	CYTb	11,320	11,340	11,360	11,380	11,400	
C. pacifica	AATTGCATTA GTTCTATCTA TTGCAATTTT AGCAATTTTA CCATTCTATC ATTTAAGCAA GTTCCGAGGT ATTCAATTCT ACCCTATTAA TCAAATTTTA						11397
Translation ORF/CDS	A L V L S A L A L P F Y H L S K F R G Q F Y P N Q L						
	CYTb						
Mega1	 T						11399
Translation ORF/CDS	A L V L S A L A L P F Y H L S K F R G Q F Y P N Q L						
	CYTb						
Mega2	 T						11398
Translation ORF/CDS							
Consensus	AATTGCATTA GTT-TATCTA TTGCAATTTT AGCAATTTTA CCATTCTATC ATTTAAG-AA GTTCCGAGGT ATTCAATTCT ACCCTATTAA TCAAATTTTA						
Conservation

	11,720	11,740	11,760	11,780	11,800	11,820	11,840	11,860	11,880	11,900	11,920	11,940	11,960	11,980	12,000
<i>C. pacifica</i>	ATGAAAAAGG	TAAAAAATCT	TTTCAAGCTA	AATACATTAA	CTTATCATAT	CGAAACCGAG	GTAAAGTTCC	ACGTACTCAA	ATAAATATAA	ATGAAACAAA					
Translation ORF/CDS	S F P	L F S	K W A L	Y M L	K D Y	R F R P	L T G	R V W	I F M F	S V F					
															ND1
Mega1															C.
Translation ORF/CDS	S F P	L F S	K W A L	Y M L	K D Y	R F R P	L T G	R V W	I F M F	S V F					
															ND1
Mega2															T
Translation ORF/CDS	S F P	L F S	K W A L	Y M L	K D Y	R F R P	L T G	R V W	I F M F	S V F					
Consensus	ATGAAAAAGG	TAAAAAATCT	TTTCAAGCTA	AATACATTAA	CTTATCATAT	CGAAA-CGAG	GTAAAGTTCC	ACGTACTCAA	ATAAATATAA	ATGA-ACAAA					
Conservation															
<i>C. pacifica</i>	TGTTAACTTA	ATATAAAAAA	AGAAAGAAAA	TACATCTCTT	CCTAAGAATA	TTACACAAAA	TAATATACTT	ATAAATAAAA	TACTTGCATA	TTCAGCCAAA					
Translation ORF/CDS	T L K	I Y F F	F S F	V D S	G L F M	V C F	L M S	M F L	I S A Y	E A L					
															ND1
Mega1															
Translation ORF/CDS	T L K	I Y F F	F S F	V D S	G L F M	V C F	L M S	M F L	I S A Y	E A L					
															ND1
Mega2															C.
Translation ORF/CDS	T L K	I Y F F	F S F	V D S	G L F M	V C F	L M S	M F L	I S A Y	E A L					
Consensus	TGTTAACTTA	ATATAAAAAA	AGAAAGAAAA	TACATCTCTT	CCTAAGAATA	TTACACAAAA	TAATATACTT	ATAAATAAAA	TACTTGCATA	TTCAGCCAAA					
Conservation															
<i>C. pacifica</i>	AAAATTAATG	CAAATCCTCC	TCTTCTATAT	TCTACATTAA	ATCCAGAAAC	TAATTCTGAT	TCTCCTTCCG	CAAAATCAAA	AGGTGTACGA	TTTGTITCTG					
Translation ORF/CDS	F I L A	F G G	S S Y	E V N F	G S V	L E S	E G E A	F D F	P T R	N T E A					
															ND1
Mega1															
Translation ORF/CDS	F I L A	F G G	S S Y	E V N F	G S V	L E S	E G E A	F D F	P T R	N T E A					
															ND1
Mega2															C.
Translation ORF/CDS	F I L A	F G G	S S Y	E V N F	G S V	L E S	E G E A	F D F	P T R	N T E A					
Consensus	AAAATTAATG	CAAATCCTCC	TCTTCTATAT	TCTACATTAA	ATCCAGAAAC	TAATTCTGAT	TC-CCTTCCG	CAAAATCAAA	AGGTGTACGA	TTTGTITCTG					
Conservation															

		12,020		12,040		12,060		12,080		ND1																									
C. pacifica	CTAAAGAAAT	ACTAAATCAA	ACTAAAGCTA	TAGGAAATAT	AATAAATAAA	AATCAAATAA	ATATTTGATA	CTTATAAAAT	ATTAATATAT	TATAACCCCC	12097																								
Translation ORF/CDS	L	S	I	S	F	W	V	L	A	M	P	F	M	I	F	L	F	W	I	F	M	Q	Y	K	Y	F	M	L	M	N	Y	G	G	ND1	
Mega1										A																								12099	
Translation ORF/CDS	L	S	I	S	F	W	V	L	S	M	P	F	M	I	F	L	F	W	I	F	M	Q	Y	K	Y	F	M	L	M	N	Y	G	G	ND1	
Mega2																																		12098	
Translation ORF/CDS	L	S	I	S	F	W	V	L	A	M	P	F	M	I	F	L	F	W	I	F	M	Q	Y	K	Y	F	M	L	M	N	Y	G	G	ND1	
Consensus	CTAAAGAAAT	ACTAAATCAA	ACTAAAG-TA	TAGGAAATAT	AATAAATAAA	AATCAAATAA	ATATTTGATA	CTTATAAAAT	ATTAATATAT	TATAACCCCC																									
Conservation																																			
		12,120		12,140		12,160		12,180		ND1																									
C. pacifica	AATTAAGAAA	ATAAACTTA	ATAAACTAA	TGCTAAACTT	ACTTCATAAG	AAATAGTTTG	AGCAACAGCT	CGTAACCCCTC	CTAATAAAGC	ATAATTAGAA	12197																								
Translation ORF/CDS	I	L	F	I	F	S	L	L	V	L	A	L	S	V	E	Y	S	I	T	Q	A	V	A	R	L	G	G	L	L	A	Y	N	S	ND1	
Mega1																																		12199	
Translation ORF/CDS	I	L	F	I	F	S	L	L	V	L	A	L	S	V	E	Y	S	I	T	Q	A	V	A	R	L	G	G	L	L	A	Y	N	S	ND1	
Mega2																																		12198	
Translation ORF/CDS	I	L	F	I	F	S	L	L	V	L	A	L	S	V	E	Y	S	I	T	Q	A	V	A	R	L	G	G	L	L	A	Y	N	S	ND1	
Consensus	AATTAAGAAA	ATAAACTTA	ATAAACTAA	TGCTAAACTT	ACTTCATAAG	AAATAGTTTG	AGCAACAGCT	CGTAACCCCTC	CTAATAAAGC	ATAATTAGAA																									
Conservation																																			
		12,220		12,240		12,260		12,280		ND1																									
C. pacifica	TTAGAAGATC	AACCAGCAAT	TATTACTGTA	TAAACCCCCA	AACTAGTACA	ACATAAGAAA	AACAATAAAC	CTAAATTAAA	AGAAAAAAGC	TTTACAAATA	12297																								
Translation ORF/CDS	N	S	S	W	G	A	I	M	V	T	Y	V	G	L	S	T	C	C	L	F	F	L	L	G	L	N	F	S	F	L	K	V	F	M	ND1
Mega1																																			12299
Translation ORF/CDS	N	S	S	W	G	A	I	M	V	T	Y	V	G	L	S	T	C	C	L	F	F	L	L	G	L	N	F	S	F	L	K	V	F	M	ND1
Mega2																																			12298
Translation ORF/CDS	N	S	S	W	G	A	I	M	V	T	Y	V	G	L	S	T	C	C	L	F	F	L	L	G	L	N	F	S	F	L	K	V	F	M	ND1
Consensus	TTAGAAGATC	AACCAGCAAT	TATTACTGTA	TAAACCCCCA	AACTAGTACA	ACATAAGAAA	AACAATAAAC	CTAAATTAAA	AGAAAAAAGC	TTTACAAATA																									
Conservation

		12,920	12,940	12,960	12,980	12,996	
C. pacifica	Translation ORF/CDS	TTACGCTGTT	ATCCCTAAAG	TAACCTATTT	TTTAAATCGC	TAATAACGGA	TCAATTATTC
						ATAAATTAAT	GATTTAAACA
						ATTAAAAAGTT	TATCAAATTT
							12996
							rmL
Mega1	Translation ORF/CDS						12999
							rmL
Mega2	Translation ORF/CDS						12998
							Consensus
	Conservation	TTACGCTGTT	ATCCCTAAAG	TAACCTATTT	TTTAAATCGC	TAATAACGGA	TCAATTATTC
						ATAAATTAAT	GATTTAAACA
						ATTAAAAAGTT	TATCAAATTT
							13096
							rmL
C. pacifica	Translation ORF/CDS	TAATATCACC	CCAATAAAAT	ATTATCAATT	ATTATAACAA	AAAAAATCTA	TAAAATTATA
						ATAACCTAAA	TATAAAGATT
						TATAGGGTCT	TCTCGTCTTT
							13096
							rmL
Mega1	Translation ORF/CDS						13099
							rmL
Mega2	Translation ORF/CDS						13098
							Consensus
	Conservation	TAATATCACC	CCAATAAAAT	ATTATCAATT	ATTATAACAA	AAAAAATCTA	TAAAATTATA
						ATAACCTAAA	TATAAAGATT
						TATAGGGTCT	TCTCGTCTTT
							13196
							rmL
C. pacifica	Translation ORF/CDS	TAAATATATT	TTAGCTTTTT	GACTAAAAAA	TAAAAATCTA	TTATAAATTT	TTATGAAACA
						GTTAATATCT	CGTCCAACCA
						TTCATACCAG	CCTTCAATTA
							13196
							rmL
Mega1	Translation ORF/CDS						13199
							rmL
Mega2	Translation ORF/CDS						13198
							Consensus
	Conservation	TAAATATATT	TTAGCTTTTT	GACTAAAAAA	TAAAAATCTA	TTATAAATTT	TTATGAAACA
						GTTAATATCT	CGTCCAACCA
						TTCATACCAG	CCTTCAATTA

		13,220	13,240	13,260	13,280	13,300	
<i>C. pacifica</i>	Translation ORF/CDS	AAAGACTAAT	GATTATGCTA	CCTTTGCACA	GTTAGTATAC	TGCGGCCATT	TAAAAATTAT TCAGTGGGCA GGATAGACTT TAAAAAAAAT TCAAAAAGAC 13296
							rmL
Mega1	Translation ORF/CDS						C..... 13299
							rmL
Mega2	Translation ORF/CDS						C..... 13298
	Consensus	AAAGACTAAT	GATTATGCTA	CCTTTGCACA	GTTAGTATAC	TGCGGCCATT	TAAAAATTAT TCAGTGGGCA GGATAGACTT TAAAAAAAAT TCAAAAAGAC
	Conservation						
		13,320	13,340	13,360	13,380	13,400	
<i>C. pacifica</i>	Translation ORF/CDS	ATGTTTTTGT	TAAACAGGCG	AACATTATTT	TTGCCGAGTT	CTTTATAAAA	ACTTTTCATT TATATTTATT TATACAATAC AATATACTAA TTTTATCATT 13396
							rmL
Mega1	Translation ORF/CDS						 13399
							rmL
Mega2	Translation ORF/CDS						 13398
	Consensus	ATGTTTTTGT	TAAACAGGCG	AACATTATTT	TTGCCGAGTT	CTTTATAAAA	ACTTTTCATT TATATTTATT TATACAATAC AATATACTAA TTTTATCATT
	Conservation						
		13,420	13,440	13,460	13,480	13,500	
<i>C. pacifica</i>	Translation ORF/CDS	ATTTATTTAT	TTTTATAATT	AAAATAAATA	TTTCAATAAA	AACTTAAAA	TTTAATAAAT AATAATTATT ATAAAAATAA TTATAACACT TTTATTAATA 13496
							rmL
Mega1	Translation ORF/CDS						 13499
							rmL
Mega2	Translation ORF/CDS						 13498
	Consensus	ATTTATTTAT	TTTTATAATT	AAAATAAATA	TTTCAATAAA	AACTTAAAA	TTTAATAAAT AATAATTATT ATAAAAATAA TTATAACACT TTTATTAATA
	Conservation

		14,120	14,140	14,160	14,180	14,200	14,220	14,240	14,260	14,280	14,300	14,320	14,340	14,360	14,380	14,400	14,420	14,440	14,460	14,480	14,500	14,520	14,540	14,560	14,580	14,600	14,620	14,640	14,660	14,680	14,700	14,720	14,740	14,760	14,780	14,800	14,820	14,840	14,860	14,880	14,900	14,920	14,940	14,960	14,980	15,000	15,020	15,040	15,060	15,080	15,100	15,120	15,140	15,160	15,180	15,200	15,220	15,240	15,260	15,280	15,300	15,320	15,340	15,360	15,380	15,400	15,420	15,440	15,460	15,480	15,500	15,520	15,540	15,560	15,580	15,600	15,620	15,640	15,660	15,680	15,700	15,720	15,740	15,760	15,780	15,800	15,820	15,840	15,860	15,880	15,900	15,920	15,940	15,960	15,980	16,000	16,020	16,040	16,060	16,080	16,100	16,120	16,140	16,160	16,180	16,200	16,220	16,240	16,260	16,280	16,300	16,320	16,340	16,360	16,380	16,400	16,420	16,440	16,460	16,480	16,500	16,520	16,540	16,560	16,580	16,600	16,620	16,640	16,660	16,680	16,700	16,720	16,740	16,760	16,780	16,800	16,820	16,840	16,860	16,880	16,900	16,920	16,940	16,960	16,980	17,000	17,020	17,040	17,060	17,080	17,100	17,120	17,140	17,160	17,180	17,200	17,220	17,240	17,260	17,280	17,300	17,320	17,340	17,360	17,380	17,400	17,420	17,440	17,460	17,480	17,500	17,520	17,540	17,560	17,580	17,600	17,620	17,640	17,660	17,680	17,700	17,720	17,740	17,760	17,780	17,800	17,820	17,840	17,860	17,880	17,900	17,920	17,940	17,960	17,980	18,000	18,020	18,040	18,060	18,080	18,100	18,120	18,140	18,160	18,180	18,200	18,220	18,240	18,260	18,280	18,300	18,320	18,340	18,360	18,380	18,400	18,420	18,440	18,460	18,480	18,500	18,520	18,540	18,560	18,580	18,600	18,620	18,640	18,660	18,680	18,700	18,720	18,740	18,760	18,780	18,800	18,820	18,840	18,860	18,880	18,900	18,920	18,940	18,960	18,980	19,000	19,020	19,040	19,060	19,080	19,100	19,120	19,140	19,160	19,180	19,200	19,220	19,240	19,260	19,280	19,300	19,320	19,340	19,360	19,380	19,400	19,420	19,440	19,460	19,480	19,500	19,520	19,540	19,560	19,580	19,600	19,620	19,640	19,660	19,680	19,700	19,720	19,740	19,760	19,780	19,800	19,820	19,840	19,860	19,880	19,900	19,920	19,940	19,960	19,980	20,000	20,020	20,040	20,060	20,080	20,100	20,120	20,140	20,160	20,180	20,200	20,220	20,240	20,260	20,280	20,300	20,320	20,340	20,360	20,380	20,400	20,420	20,440	20,460	20,480	20,500	20,520	20,540	20,560	20,580	20,600	20,620	20,640	20,660	20,680	20,700	20,720	20,740	20,760	20,780	20,800	20,820	20,840	20,860	20,880	20,900	20,920	20,940	20,960	20,980	21,000	21,020	21,040	21,060	21,080	21,100	21,120	21,140	21,160	21,180	21,200	21,220	21,240	21,260	21,280	21,300	21,320	21,340	21,360	21,380	21,400	21,420	21,440	21,460	21,480	21,500	21,520	21,540	21,560	21,580	21,600	21,620	21,640	21,660	21,680	21,700	21,720	21,740	21,760	21,780	21,800	21,820	21,840	21,860	21,880	21,900	21,920	21,940	21,960	21,980	22,000	22,020	22,040	22,060	22,080	22,100	22,120	22,140	22,160	22,180	22,200	22,220	22,240	22,260	22,280	22,300	22,320	22,340	22,360	22,380	22,400	22,420	22,440	22,460	22,480	22,500	22,520	22,540	22,560	22,580	22,600	22,620	22,640	22,660	22,680	22,700	22,720	22,740	22,760	22,780	22,800	22,820	22,840	22,860	22,880	22,900	22,920	22,940	22,960	22,980	23,000	23,020	23,040	23,060	23,080	23,100	23,120	23,140	23,160	23,180	23,200	23,220	23,240	23,260	23,280	23,300	23,320	23,340	23,360	23,380	23,400	23,420	23,440	23,460	23,480	23,500	23,520	23,540	23,560	23,580	23,600	23,620	23,640	23,660	23,680	23,700	23,720	23,740	23,760	23,780	23,800	23,820	23,840	23,860	23,880	23,900	23,920	23,940	23,960	23,980	24,000	24,020	24,040	24,060	24,080	24,100	24,120	24,140	24,160	24,180	24,200	24,220	24,240	24,260	24,280	24,300	24,320	24,340	24,360	24,380	24,400	24,420	24,440	24,460	24,480	24,500	24,520	24,540	24,560	24,580	24,600	24,620	24,640	24,660	24,680	24,700	24,720	24,740	24,760	24,780	24,800	24,820	24,840	24,860	24,880	24,900	24,920	24,940	24,960	24,980	25,000	25,020	25,040	25,060	25,080	25,100	25,120	25,140	25,160	25,180	25,200	25,220	25,240	25,260	25,280	25,300	25,320	25,340	25,360	25,380	25,400	25,420	25,440	25,460	25,480	25,500	25,520	25,540	25,560	25,580	25,600	25,620	25,640	25,660	25,680	25,700	25,720	25,740	25,760	25,780	25,800	25,820	25,840	25,860	25,880	25,900	25,920	25,940	25,960	25,980	26,000	26,020	26,040	26,060	26,080	26,100	26,120	26,140	26,160	26,180	26,200	26,220	26,240	26,260	26,280	26,300	26,320	26,340	26,360	26,380	26,400	26,420	26,440	26,460	26,480	26,500	26,520	26,540	26,560	26,580	26,600	26,620	26,640	26,660	26,680	26,700	26,720	26,740	26,760	26,780	26,800	26,820	26,840	26,860	26,880	26,900	26,920	26,940	26,960	26,980	27,000	27,020	27,040	27,060	27,080	27,100	27,120	27,140	27,160	27,180	27,200	27,220	27,240	27,260	27,280	27,300	27,320	27,340	27,360	27,380	27,400	27,420	27,440	27,460	27,480	27,500	27,520	27,540	27,560	27,580	27,600	27,620	27,640	27,660	27,680	27,700	27,720	27,740	27,760	27,780	27,800	27,820	27,840	27,860	27,880	27,900	27,920	27,940	27,960	27,980	28,000	28,020	28,040	28,060	28,080	28,100	28,120	28,140	28,160	28,180	28,200	28,220	28,240	28,260	28,280	28,300	28,320	28,340	28,360	28,380	28,400	28,420	28,440	28,460	28,480	28,500	28,520	28,540	28,560	28,580	28,600	28,620	28,640	28,660	28,680	28,700	28,720	28,740	28,760	28,780	28,800	28,820	28,840	28,860	28,880	28,900	28,920	28,940	28,960	28,980	29,000	29,020	29,040	29,060	29,080	29,100	29,120	29,140	29,160	29,180	29,200	29,220	29,240	29,260	29,280	29,300	29,320	29,340	29,360	29,380	29,400	29,420	29,440	29,460	29,480	29,500	29,520	29,540	29,560	29,580	29,600	29,620	29,640	29,660	29,680	29,700	29,720	29,740	29,760	29,780	29,800	29,820	29,840	29,860	29,880	29,900	29,920	29,940	29,960	29,980	30,000	30,020	30,040	30,060	30,080	30,100	30,120	30,140	30,160	30,180	30,200	30,220	30,240	30,260	30,280	30,300	30,320	30,340	30,360	30,380	30,400	30,420	30,440	30,460	30,480	30,500	30,520	30,540	30,560	30,580	30,600	30,620	30,640	30,660	30,680	30,700	30,720	30,740	30,760	30,780	30,800	30,820	30,840	30,860	30,880	30,900	30,920	30,940	30,960	30,980	31,000	31,020	31,040	31,060	31,080	31,100	31,120	31,140	31,160	31,180	31,200	31,220	31,240	31,260	31,280	31,300	31,320	31,340	31,360	31,380	31,400	31,420	31,440	31,460	31,480	31,500	31,520	31,540	31,560	31,580	31,600	31,620	31,640	31,660	31,680	31,700	31,720	31,740	31,760	31,780	31,800	31,820	31,840	31,860	31,880	31,900	31,920	31,940	31,960	31,980	32,000	32,020	32,040	32,060	32,080	32,100	32,120	32,140	32,160	32,180	32,200	32,220	32,240	32,260	32,280	32,300	32,320	32,340	32,360	32,380	32,400	32,420	32,440	32,460	32,480	32,500	32,520	32,540	32,560	32,580	32,600	32,620	32,640	32,660	32,680	32,700	32,720	32,740	32,760	32,780	32,800	32,820	32,840	32,860	32,880	32,900	32,920	32,940	32,960	32,980	33,000	33,020	33,040	33,060	33,080	33,100	33,120	33,140	33,160	33,180	33,200	33,220	33,240	33,260	33,280	33,300	33,320	33,340	33,360	33,380	33,400	33,420	33,440	33,460	33,480	33,500	33,520	33,540	33,560	33,580	33,600	33,620	33,640	33,660	33,680	33,700	33,720	33,740	33,760	33,780	33,800	33,820	33,840	33,860	33,880	33,900	33,920	33,940	33,960	33,980	34,000	34,020	34,040	34,060	34,080	34,100	34,120	34,140	34,160	34,180	34,200	34,220	34,240	34,260	34,280	34,300	34,320	34,340	34,360	34,380	34,400	34,420	34,440	34,460	34,480	34,500	34,520	34,540	34,560	34,580	34,600	34,620	34,640	34,660	34,680	34,700	34,720	34,740	34,760	34,780	34,800	34,820	34,840	34,860	34,880	34,900	34,920	34,940	34,960	34,980	35,000	35,020	35,040	35,060	35,080	35,100	35,120	35,140	35,160	35,180	35,200	35,220	35,240	35,260	35,280	35,300	35,320	35,340	35,360	35,380	35,400	35,420	35,440	35,460	35,480	35,500	35,520	35,540	35,560	35,580	35,600	35,620	35,640	35,660	35,680	35,700	35,720	35,740	35,760	35,780	35,800	35,820	35,840	35,860	35,880	35,900	35,920	35,940	35,960	35,980	36,000	36,020	36,040	36,060	36,080	36,100	36,120	36,140	36,160	36,180	36,200	36,220	36,240	36,260	36,280	36,300	36,320	36,340	36,360	36,380	36,400	36,420	36,440	36,460	36,480	36,500	36,520	36,540
--	--	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

C. pacifica 15,820 15,840 15,860 15,880 15,900
 Translation ORF/CDS CATAATTTAA TATATTATAA ATATATGAAT AAATATATAT AAATTATATA TATATATATA TATGCAATTT TTAATTATCT AAATTTATTT TTTTTCACCTT 15824

Mega1 15267
 Translation ORF/CDS

Mega2 A.. 15698
 Translation ORF/CDS

Consensus CATAATTTAA TATATTATAA ATATATGAAT AAATATATAT AAATTATATA TATATATATA TATGCAATTT TTAATTATCT AAATTTATTT TTTTTCACCTT
 Conservation

C. pacifica 15,720 15,740 15,760 15,780 15,800
 Translation ORF/CDS TATAAACTAA TTTACTATTT ATAATTATTA ATATTTAAAT TTTATAATAT TATTTTATTT AATAAATCAC AAAAAAANNA CTATAAAATA AA----- 15716

Mega1 15273
 Translation ORF/CDS

Mega2 G. TT. ACAAAAAA 15798
 Translation ORF/CDS

Consensus TATAAACTAA TTTACTATTT ATAATTATTA ATATTTAAAT TTTATAAT-T TATTTTATTT AATAAATCAC AAAAAA-A CTATAAAATA AAACAAAAAA
 Conservation

C. pacifica 15,820
 Translation ORF/CDS 15716

Mega1 15273
 Translation ORF/CDS

Mega2 AAAAAAAAAA AAAAAACAT GGGAATCTTA CTATT 15833
 Translation ORF/CDS

Consensus AAAAAAAAAA AAAAAACAT GGGAATCTTA CTATT
 Conservation

APPENDIX D

PROPOSED SECONDARY STRUCTURES OF RNA MOLECULES

Secondary structure: $\Delta G = -9.0\text{kcal/mol}$

A

Secondary structure: $\Delta G = -3.5\text{kcal/mol}$

E

Secondary structure: $\Delta G = -10.9\text{kcal/mol}$

C

Secondary structure: $\Delta G = -7.9\text{kcal/mol}$

F

Secondary structure: $\Delta G = -5.8\text{kcal/mol}$

D

G¹

H¹

Secondary structure: $\Delta G = -11.8 \text{ kcal/mol}$

Secondary structure: $\Delta G = -12.3 \text{ kcal/mol}$

I

Secondary structure: $\Delta G = -11.3 \text{ kcal/mol}$

Secondary structure: $\Delta G = -9.8 \text{ kcal/mol}$

K

Secondary structure: $\Delta G = -9.9 \text{ kcal/mol}$

L1²

L2²

M

Secondary structure: $\Delta G = -11.1 \text{ kcal/mol}$

N

Secondary structure: $\Delta G = -8.5 \text{ kcal/mol}$

R

Secondary structure: $\Delta G = -9.3 \text{ kcal/mol}$

P

Secondary structure: $\Delta G = -13.0 \text{ kcal/mol}$

S1²

Secondary structure: $\Delta G = -7.8 \text{ kcal/mol}$

Q

Secondary structure: $\Delta G = -11.3 \text{ kcal/mol}$

S2²

T¹

Secondary structure: $\Delta G = -11.9 \text{ kcal/mol}$

Secondary structure: $\Delta G = -8.3 \text{ kcal/mol}$

V

W¹

Y

tRNAs are shown in alphabetical order. A: Alanine. C: Cysteine. D: Aspartate. E: Glutamate. F: Phenylalanine. G: Glycine¹. H: Histidine¹. I: Isoleucine. K: Lysine. L: Leucine². M: Methionine. N: Asparagine. P: Proline. Q: Glutamine. R: Arginine. S: Serine². T: Threonine¹. V: Valine. W: Tryptophan¹. Y: Tyrosine.

¹: CLC could not determine an acceptable secondary structure. Images from MITOS (<http://mitos.bioinf.uni-leipzig.de/index.py>) but they lack a reportable folding energy. All other structures produced in CLC Genomics Workbench.

²: Leucine and Serine have two tRNAs each. These are indicated by Arabic numerals and differ in their anticodon.

16s rRNA for *Chrysomya pacifica*

Source: CLC Genomics Workbench

$\Delta G = -173.4 \text{ kcal/mol}$

16s rRNA *Chrysomya pacifica*

Source: MITOS (<http://mitos.bioinf.uni-leipzig.de/index.py>)

16s rRNA *Chrysomya pacifica*
source: LocARNA
(<http://rna.informatik.uni-freiburg.de>)

12s rRNA *Chrysomya pacifica*

Source: CLC Genomics Workbench

$\Delta G = -114.5 \text{ kcal/mol}$

12s rRNA *Chrysomya pacifica*

Source: MITOS (<http://mitos.bioinf.uni-leipzig.de/index.py>)

16s rRNA *Chrysomya pacifica*
source: LocARNA
(<http://rna.informatik.uni-freiburg.de>)

Appendix E

ELECTROPHEROGRAMS FOR STR PRIMER 40-15840

LE 07/015

Sample Name	Size³	Size³
Liz600¹		
S01	92.27	92.27
S02	85.67	95.39
S03	85.59	92.30
S04	92.13	
S05	82.51	85.71
S06	92.50	92.50
S07	92.42	95.59
S08	82.53	85.63
S09²		
S10	86.25	92.34
S11	85.68	95.39
S12	85.59	92.30
S13	85.71	92.30
S14	92.33	92.33
S15	82.41	85.62
S16	85.68	95.39
S17	85.74	92.49
S18²		
15_21	85.78	95.45
S19	92.32	92.32
S20	92.46	95.55
S21	92.47	92.47
S22	85.78	92.45
S23	82.68	85.82
S24	85.72	92.43
S25	92.49	92.49
S26²		

¹:LIZ600 is the capillary electrophoresis blank.

²:S09, S18, and S26 are PCR blanks.

³:Size in base pairs as determined by Genemapper and ABI 3130 Collection Software.