Alternative splicing of DNA polymerase beta in vertebrates

by

Emma Bondy-Chorney, B.Sc.

Submitted in partial fulfillment
of the requirements for the degree of
Master of Science

Department of Biological Sciences, Brock University
St. Catharines, Ontario

Abstract

Alternative splicing (AS) is the predominant mechanism responsible for increasing eukaryotic transcriptome and proteome complexity. In this phenomenon, numerous mRNA transcripts are produced from a single pre-mRNA sequence. AS is reported to occur in 95% of human multi-exon genes; one specific gene that undergoes AS is DNA polymerase beta (*POLB*). *POLB* is the main DNA repair gene which performs short patch base excision repair (BER). In primate untransformed primary fibroblast cell lines, it was determined that the splice variant (SV) frequency of *POLB* correlates positively with species lifespan. To date, AS patterns of *POLB* have only been examined in mammals primarily through the use of cell lines. However, little attention has been devoted to investigating if such a relationship exists in non-mammals and whether cell lines reflect what is observed in vertebrate tissues. This idea was explored through cloning and characterization of 1,214 *POLB* transcripts from four non-mammalian species (*Gallus gallus domesticus, Larus glaucescens, Xenopus laevis,* and *Pogona vitticeps*) and two mammalian species (*Sylvilagus floridanus* and *Homo sapiens*) in two tissue types, liver and brain. *POLB* SV frequency occurred at low frequencies, $\leq$ 3.2%, in non-mammalian tissues relative to mammalian ($\geq$20%). The highest *POLB* SV frequency was found in *H. sapiens* liver and brain tissues, occurring at 65.4% and 91.7%, respectively. Tissue specific AS of *POLB* was observed in *L. glaucescens, P. vitticeps*, and *H. sapiens,* but not *G. gallus domesticus*, *X. laevis* and *S. floridanus*.

The AS patterns of a second gene, transient receptor potential cation channel subfamily V member 1 (*TRPV1*), were compared to those of *POLB* in liver and brain

tissues of *G. gallus domesticus*, *X. laevis* and *H. sapiens*. This comparison was performed to investigate if any changes (either increase or decrease) observed in the AS of *POLB* were gene specific or if they were tissue specific, in which case similar changes in AS would be seen in *POLB* and *TRPV1*. Analysis did not reveal an increase or decrease in both the AS of *POLB* and *TRPV1* in either the liver or brain tissues of *G. gallus domesticus* and *H. sapiens*. This result suggested that the AS patterns of *POLB* were not influenced by tissue specific rates of AS. Interestingly, an increase in the AS of both genes was only observed in *X. laevis* brain tissue. This result suggests that AS in general may be increased in the *X. laevis* brain as compared to liver tissue.

No positive correlation between *POLB* SV frequency and species lifespan was found in non-mammalian tissues. The AS patterns of *POLB* in human primary untransformed fibroblast cell lines were representative of those seen in human liver tissue but not in brain tissue. Altogether, the AS patterns of *POLB* from vertebrate tissues and primate cell lines revealed a positive correlation between *POLB* SV frequency and lifespan in mammals, but not in non-mammals. It appears that this positive correlation does not exist in vertebrate species as a whole.

Acknowledgments

I would first like to thank science. I am so grateful that I live in a time when I am able to study and understand the world around me. For as long as I can remember I have loved science and I am so happy that I am able to spend my whole life continuously being amazed. I believe that understanding the mysteries of the universe does nothing to take away from the awe and wonder of it.

Thank you to my supervisor, Dr. Adonis Skandalis. I will always be grateful for your continuous support and everything you have taught me over the years. Your voice will always be there to remind me that this isn't "space science" and to remember that the devil is in the details. To my committee members, Dr. Ping Liang and Dr. Yousef Haj-Ahmad, the help you have given me with this work over the past two years is much appreciated, thank you. I could not have done this thesis without the help of Dr. Jeff Stuart, Dr. Robert Carlone and Dr. Glenn Tattersall and all the students in their labs for providing tissues samples and really good discussions.

I owe so much to my friends here at Brock and would like to give a special thank you to Amanda Lepp, Jess Vickruck, Chris Course, Carrie Snajder and Taylor Dawson. You guys took the time and patience to help me and be my friends over the years, and that couldn't have always been easy!

To my family, I cannot explain how much your love and support has meant to me during this time. My amazingly strong sister Liz who has always been there telling me I could do this, my incredibly patient father Richard whose confidence in me kept me going and finally my mother Jan, whose strength and love means the world to me and has shown me I can do anything I put my mind to. Lastly, I want to thank Stacy for always being there for me and listening to my crazy rants about RNA and frog brains. You are truly an amazing man and I am so lucky to have you in my life.

Table of Contents

Table of abbreviations

ACUC	Animal Care and Use Committee
AS	Alternative splicing
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
cDNA	Complementary deoxyribonucleic acid
CDS1	Cytidylyltransferase 1
CT	Canonical transcript
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Triphosphates
Dscam	Down syndrome cell adhesion molecule
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraaceticacid
ESC	Embryonic stem cells
ESEs	Exonic splicing enhancers
ESSs	Exonic splicing silencers
ESTs	Expressed sequence tags
FOXP1	Forkhead box protein P1 gene
hnRNPs	heterogeneous nuclear ribonuclear proteins
ISEs	Intronic splicing enhancers
ISSs	Intronic splicing silencers
ITPG	Isopropyl-β-D-thio-galactoside
LB	Luria Broth
MBNL	Musclebind-like
mRNA	messenger ribonucleic acid
NMD	Nonsense mediated decay
p53	Tumour protein 53
PCR	Polymerase chain reaction
RNP	Ribonucleoprotein
POLB	DNA polymerase beta
PPT	Polypyrimidine tract
PTB	Polypyrimdine tract-binding
PTC	Premature termination codons
RBFOX	RNA-binding splicing factor
ROS	Reactive oxygen species
RNA	Ribonucleic acid
rpL3	Ribosomal protein L3
SF2	Splicing factor 2
SHOX	Short stature homeobox
snRNAs	small nuclear RNAs
SSs	Splice sites
STREX	Stress axis-regulated exon
SV	Splice variant
TAE	Tris-Acetic acid
TM	Transmembrane

TRPV1	Transient receptor potential cation channel subfamily V member 1
UV	Ultraviolet
WT	Wild type
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

List of tables

List of figures

General introduction and thesis overview

The central dogma of molecular biology states that genetic information flows from deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) to proteins (Crick 1970). Eukaryotic gene expression requires a regulatory stage called RNA splicing which involves the removal of introns and the ligation of exons to form messenger RNA (mRNA) (Darnell 1978). The mRNA then is processed through a step called translation to produce proteins. In the human genome, there are over 100,000 predicted proteins synthesized from an estimated 24,000 protein coding genes (Modrek et al., 2002). The phenomenon of alternative splicing is responsible for generating such a diverse array of proteins from only a limited number of mRNAs (Chow et al., 1977). Alternative splicing can produce many different forms of transcripts from one gene, called splice variants, through the inclusion of different exons in the mRNA. Remarkably, in humans 95% of multi-exon genes undergo alternative splicing (Wang et al., 2008). In addition to generating protein diversity, the process of alternative splicing has been linked to gene regulation, gene expression, disease, and life history traits, such as lifespan (Taliaferro et al., 2011; Skandalis et al., 2010; Pan et al., 2008; Su et al., 2004; Yeo, et al., 2004).

A positive correlation between DNA polymerase beta (*POLB*) splice variant frequency and the lifespan of an organism was reported in primate untransformed fibroblast cell lines (Skandalis et al., 2010). Interestingly, *POLB* is a DNA repair gene (Singhal et al., 1993) and the *POLB* splice variants generated as a result of alternative splicing are proposed to be unproductive; in other words, they do not form functional proteins (Skandalis et al., 2004). Why would the production of non-productive splice variants of a DNA repair gene correlate positively with lifespan? Although the alternative

splicing patterns of *POLB* in mammals are well documented, with data collections compiled largely from human cell lines and several tissues (Skandalis et al., 2010; Simonelli et al., 2009; Skandalis et al., 2004; Thompson et al., 2002), little to no research has been focused on the alternative splicing patterns of *POLB* in non-mammals. Does a positive correlation between the alternative SV frequency of *POLB* and lifespan suggested in primates exist in non-primate mammals and in general in vertebrates? This thesis investigates said question by examining the alternative splicing of *POLB* in liver and brain tissues of six vertebrate species, four non-mammals and two mammals.

The first chapter is a literature review introducing RNA splicing, alternative splicing, *POLB*, the alternative splicing of *POLB* and the concept of lifespan as it relates to this work. This chapter provides the background necessary to investigate the correlation between alternative splicing of *POLB* and lifespan in vertebrates examined in Chapters Two and Three.

Various questions must be addressed before the subject of alternative splicing of *POLB* and lifespan in non-mammals can be discussed. Does *POLB* undergo alternative splicing in non-mammals? If so, what are the *POLB* splice variants types and their frequencies? Is the alternative splicing pattern of *POLB* tissue specific in these species? Could these patterns be influenced by an overall increase in alternative splicing in a specific tissue type? How do the alternative splicing patterns of *POLB* compare to other genes in the same tissues? Does a positive correlation between *POLB* splice variant frequency and lifespan exist in non-mammals? Chapter Two will address these questions through the characterization of over 840 *POLB* in non-mammalian tissues.

Chapter Three compares the alternative splicing of *POLB* in mammalian tissues to those described in primate untransformed fibroblast cell lines (Skandalis et al., 2010). This comparison is essential to investigate whether the alternative splicing of *POLB* in tissues reflects that observed in cell lines. A detailed description of the alternative SV frequency of *POLB* and the types of splice variants in human, rabbit liver and brain tissues are documented for the first time. Finally, in conjunction with the alternative splicing patterns of *POLB* in mammals, non-mammals and cell lines, the possibility of a positive correlation between *POLB* splice variant frequency and lifespan in vertebrates will be explored.

CHAPTER ONE: Literature review

RNA splicing

Eukaryotic genes are composed of introns and exons. Introns include any sequence removed from the final RNA transcript, referred to as the mature messenger RNA (mRNA). Exons are defined as any sequence left in the mRNA. Before the RNA, termed as precursor-mRNA (pre-mRNA) can be used to make proteins, intronic sequences must be cut out. Once the introns are removed and exons are ligated together, mRNA is formed and protein synthesis takes place. The process of intron removal is termed RNA splicing and is carried out by the spliceosome (Zhou et al., 2002).

The spliceosome functions to recognize introns and exons in order to form the correct mRNA. The major eukaryotic spliceosome is a multimegadalton ribonucleoprotein (RNP) complex comprised of five small nuclear RNAs (snRNAs) (U1, U2, U4, U5 and U6) and numerous proteins (Wahl et al., 2009). These five snRNAs assemble with their associated proteins to form small nuclear ribonucleoproteins (snRNPs). Once assembled the spliceosome must be able to recognize introns and exons. This recognition is enabled by four conserved signals, 5'splice sites (5'SS), 3' splice sites (3'SS), the branch site (A), and the polypyrimidine tract (PPT) (Figure 1.1) (reviewed by Keren et al., 2010). The 5' SS is also called the donor site and is located at the 5' end of the intron. The conserved nature of this site, comprosed of a guanine (G) followed by a uracil (U), facilitates recognition. The 3' SS, also called the acceptor site, is located at the 3' end of the intron and is denoted by an adenine (A) followed by a guanine (G). The PPT is approximately 15-20 nucleotides long and rich with pyrimidine nucleotides, usually uracils. This sequence is located 5-40 base pairs before the 3' end of the intron (Wagner

et al., 2001). Upstream from the PPT is the branch site which is an A residue. The spliceosome catalyzes the two transesterification steps required for proper RNA splicing. In the first step the 2'hydroxyl group of the branch site attacks the phosphate at the 5'SS. Ligation of the 5'end of the intron to the branch site results in a detached 5' exon and an intron-3' exon fragment. Next, the phosphate at the 3' end of the intron is attacked by the 3'hydroxyl of the detracted exon. The result of this attack is the ligation of the two exons and removal of the intron which was in between (Black, 2003).

Splicing can be enhanced or silenced through the binding of auxiliary proteins to additional sites within exons and introns. Examples of such proteins include serine/arginine (SR) proteins, heterogeneous nuclear ribonuclear proteins (hnRNPs), and polypyrimdine tract-binding (PTB) proteins. These proteins can recognize binding sites consisting of short degenerate sequences located within exons and introns. There are four major sites, two which result in splicing enhancing and two for splicing silencing. Exons can contain exonic splicing enhancers (ESEs) or exonic splicing silencers (ESSs) and introns can contain intronic splicing enhancers (ISEs) or intronic splicing silencers (ISSs) (Keren et al., 2010). ESEs are commonly located in constitutive exons as opposed to alternatively spliced exons (Black 2003). Both of these types of exons can be involved in a process called alternative splicing (AS).

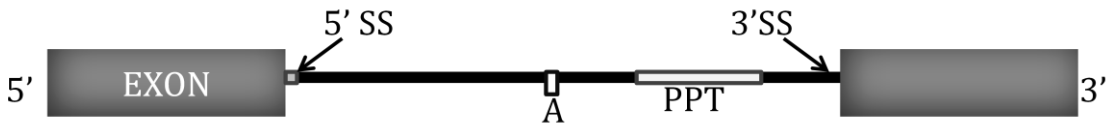

Figure 1.1 Splice signals within pre-mRNA transcripts.

Alternative Splicing

Almost immediately after the discovery of RNA splicing, it was determined that numerous mRNA transcripts could be produced from one pre-mRNA sequence (Roger and Sharp, 1977). This phenomenon was termed alternative splicing (AS) and it's importance was not immediately recognized. The significance of AS was initially underestimated, as it was thought to only affect a few human genes such as mammalian immunoglobuins and T-cell receptors. AS is now considered to be one of the major mechanisms for generating protein diversity, and is observed in metazoan, fungal, and plant genomes (Artamonova et al, 2007). Perhaps one of the most famous examples of AS can be seen in the *Drosophila melanogaster* gene Down syndrome cell adhesion molecule (*Dscam*), which can generate 38,016 distinct mRNA isoforms, a number far in excess of the total number of genes (~14,500) in the organism (Graveley, 2004). AS of human genes is extremely common (Pan et al., 2008). In fact, 95% of all human multi-exon genes are alternatively spliced (Wang et al., 2008). There are several ways a pre-mRNA sequence can be alternatively spliced (as reviewed by Modrek et al., 2002) and there are four main subgroups of AS events outlined below (Figure 1.2).

The four main subgroups of AS events are exon skipping, alternative 3'SS selection, alternative 5'SS selection and intron retention. The most common of these AS events seen in mammals is exon skipping (Sugnet et al., 2004). Exon skipping occurs when an exon, termed a cassette exon, and its flanking introns are spliced out and not included in the mRNA transcript. A cassette exon is either retained in the mRNA or alternatively used during AS (Graveley 2001). A plausible cause of this AS event is the inhibition of an exon inclusion signal located within the cassette exon itself (Van Ommen

et al., 2011). Alternative 3'SS and 5'SS selection is observed when two or more splice sites are recognized at one of the ends of an exon. This type of AS event occurs when two competing splice sites are present. The effect of competing 5'SS selection was demonstrated with a splicing factor call splicing factor 2 (SF2). SF2 is a 33 kDa protein that binds nonspecifically to RNA. Different concentrations of SF2 have a profound effect on a model pre-mRNA substrate that contains two competing 5'SS and only one 3'SS. If SF2 is in low concentration the outside 5'SS is used. In contrast, when SF2 is in high concentrations the inside 5'SS is used (Kranier et al., 1990). The competition among multiple splice sites has been linked to splicing silencers and enhancers (Nilsen et al., 2010). Furthermore, intron retention is the fourth and rarest type of AS event seen in vertebrates. It is characterized by an intron remaining in the mRNA transcript. By convention, when an in intron is left in the mRNA it is referred to as an exon. A recent example of intron retention was found in the stress axis-regulated exon (STREX)-containing a calcium-activated big potassium channel gene. In rat hippocampal cells it was discovered that splice variant expression was regulated by intron retention (Bell et al., 2010). In addition to these four major subgroups of AS events, there are less frequent types including mutually exclusive exons, alternative promoter uses and alternative polyadenylation (Keren et al., 2010). There are several models to describe how these AS events are regulated which are outlined here (Figure 1.2).

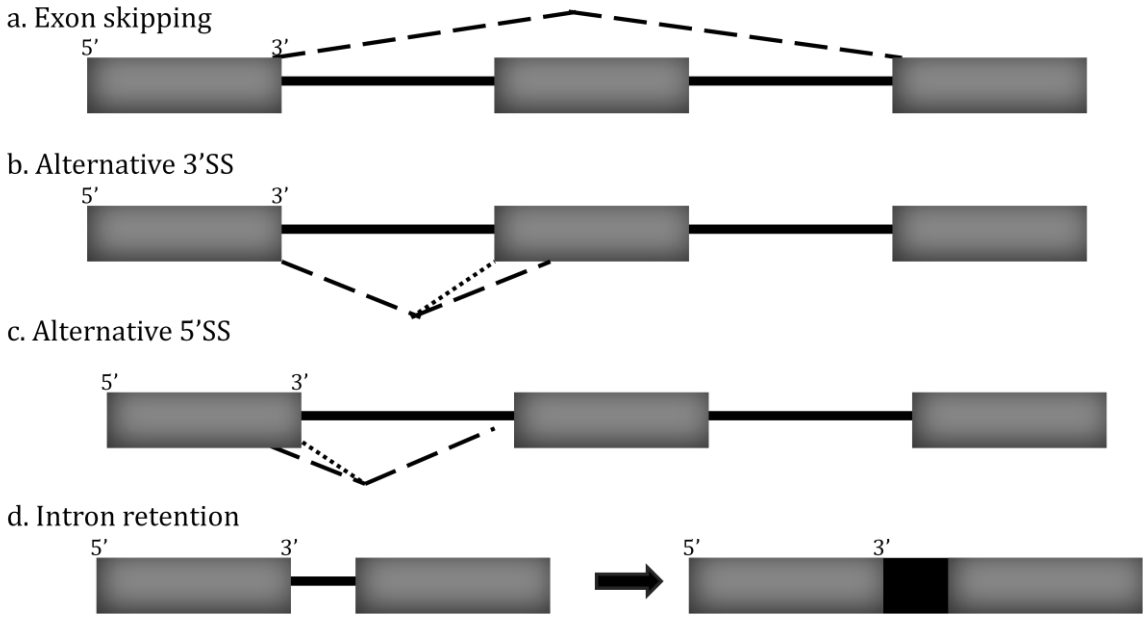

Figure 1.2 Types of AS events. The four main subgroups of AS that are seen in vertebrates. Grey boxes represent exons and black lines/boxes represent introns. Dashed lines show the splicing pattern observed in the respective type and dotted lines demonstrate the predicted splice site.

Regulation of AS

Several factors affect splicing patterns (Mata et al., 2011; Luco et al., 2001) including transcription rate, core-splicing-machinery, intron size and competition between splice sites (Nilsen et al., 2010).

Three of the major models of how AS is regulated will be presented here. The first model presented here involves the kinetic effect of splicing silencers on AS and explains regulation in terms of SS competition (Yu et al., 2008). As previously discussed, this model applies to mRNA transcripts in which an exon contains two or more splice sites. In cases where no ESS is present the appropriate spliceosome components will bind to the nearest 5'SS of the two SSs present in respective the exon. Alternatively, if the silencer is present in the exon the rate at which the first 5'SS is used is decreased

promoting the splicing of the distal 5' SS. When there is no competing 5' splice site upstream the presence of a silencer does not affect the rate of splicing (Yu et al., 2008).

Another model of AS regulation describes the impact of interplay between transcription elongation rate, chromatin structure and histone modification on AS (Alló et al., 2009). Take for example a multiple exon gene that contains both constitutively transcribed exons and one alternative exon. The constitutive exons are packaged into nucleosomes which always contain a histone that is trimethylated at the lysine residue 36 (Alló et al., 2009). Nilsen et al. (2010) proposed that when the RNA polymerase and the spliceosome transcribe a gene it travels over the alternative exon very quickly and that the exon does not become attached to the RNA complex and is not included in the mature mRNA. When the alternative exon is packaged into a nucleosome containing a histone that is trimethylated at the lysine residue 36 the processivity of the RNA polymerase is slowed. The alternative exon is captured and included in the mRNA transcript (Nilsen et al., 2010). This model agrees with the results found in a study by Alló, M. et al. (2009) in which an induced change in chromatin structure led to a reduction in transcription elongation complex ability and changes in AS patterns. Nilsen et al. (2010) explain that this model requires more investigation with particular regards to whether transcription rates would differ between tissues with high or low levels of AS using this model.

Finally, the third model for AS regulation is the Kinetic model of the mutually exclusive splicing of pre-mRNA (Neves et al., 2004). This model has been investigated primarily in the Dscam gene of *Drosophila,* previously mentioned to encode for more than 38,000 mRNA isoforms. This gene contains 48 alternative exonic sequences and determination of which exons are spliced depends on competition between RNA

secondary structures for a single docking site, one site selector sequence in each of the 48 exons and various other proteins (Nilsen et al., 2010; Schmucker et al., 2000).

In addition to these AS regulatory mechanisms nonsense mediated decay (NMD) can also play an important role in AS (Lewis et al., 2003). NMD is a quality control mechanism which functions to selectively degrade mRNAs that have premature termination codons (PTCs) in the transcript in eukaryotic cells (Hwang et al., 2011). This mechanism prevents truncated proteins from being translated, which can have negative effects (Chang et al., 2007). In humans it was found that one-third of splice variants contain PTCs. Although all of these transcripts theoretically should be subject to NMD this is not the case. There are numerous splice variants known to regulate protein expression that are not subject to NMD (Lewis et al., 2003). A recent study showed that NMD coupled with AS regulates the expression of the human short stature homeobox (*SHOX*) gene. AS of the *SHOX* gene causes an inclusion of the alternative exon 2a. This exon 2a transcript leads to NMD in turn altering the expression of *SHOX* in a tissue and time specific manner (Durand et al., 2011). As is the case with *SHOX*, it is common to find genes that are alternatively spliced in a tissue specific manner (Boue et al., 2006).

Tissue specificity and alternative splicing

Numerous AS events have been shown to be tissue specific in a variety of vertebrate species (Venables et al., 2012; Su et al., 2004; Yeo, et al., 2004). Tissue specific AS refers to both the proportion of alternative transcripts and the types of alternative transcripts produced. In some genes tissue specific AS is suggested to regulate gene expression by generating tissue specific transcripts, altering gene function, co-regulating other genes, influencing differentiation of tissues, and controlling fundamental

mechanical properties of certain tissues (Taliaferro et al., 2011, Pan et al., 2008, Su et al., 2004, Yeo, et al., 2004).

Tissue specific AS can be regulated by proteins which are themselves tissue specific (Kim et al., 2011). One example of this is the mammalian Fox 1 family (Fox) of proteins, which consist of Fox-1, Fox-2 and Fox-3. These proteins contain almost identical RNA recognition motifs that bind to the RNA element 5'-UGCAUG-3' and regulate alternative splicing. Fox-1 is only expressed in brain and striated muscles whereas Fox-2 is expressed in numerous tissues and Fox-3 produced exclusively in neural tissues. Not only do these different Fox proteins regulate AS of other genes in their respective tissues but different isoforms of the proteins have also been shown to have unique subcellular expression and an effect on splicing (Kim et al., 2011).

Tissue specific AS has been best described in mammals, most notably in humans and mouse transcriptomes (Su et al., 2003). Although recent studies investigating this phenomenon have been extensive, the available data demonstrate differences in tissue specific AS patterns. These differences are largely due to the method used to examine tissue specific patterns of AS. For example, one study analyzed splicing patterns in humans using expressed sequence tags (ESTs) gathered from cDNA libraries from different tissues (Yeo et al., 2004). This study gathered the data based on distinguishing between the different AS events exon skipping and alternative 5’SS or 3’ SS using a large scale computational analysis from approximately two dozen human tissues (Yeo et al., 2004). This analysis did not consider intron retention AS events which may have led to an underestimation of AS. Pan et al. (2008) current mRNA sequence datasets consisting of 17-32 million 32-nucleotide-long reads in seven normal human tissues to investigate

tissue specific AS patterns. In contrast to the methods employed in the study by Yeo et al. (2004), Pan et al. (2008) based their measurements for alternative splicing levels on splice junction sequences that represent known and hypothetical splicing events. Although these studies use different methods of investigation they have provided valuable information about AS tissue specificity. It is difficult to maintain consistency throughout these types of studies for various reasons. For example, when determining AS levels in tissues using ESTs and mathematical models certain types of splicing events are ignored or hypothetical splicing events are included. These factors could result in an under or over estimation of the actual AS levels in a tissue. Additionally, utilization of large databases, such as GenBank, only allow for the splicing events which have been recorded to be included in large scale studies. This could cause underestimation of AS levels (Koralewski et al., 2010).

In humans, liver and brain tissues showed the highest levels of AS (Pan et al., 2008; Yeo et al., 2004). Yeo et al. (2004) reported these levels to be 42% in the liver and 39% in the brain. This study showed other tissues were determined to have a proportion of alternative splice genes at 35% (testis), 28% (lung), 27% (kidney), and 26% (placenta). The lowest levels of AS were found to be in the uterus and muscle at 15% for both tissues (Yeo et al., 2004). This data was gathered using ESTs derived from cDNA libraries from described tissue types. The splicing events considered in this study were exon skipping and alternative 5' / 3' SS while intron retention was not considered (Yeo et al., 2004).

Tissue specific exon skipping and SS selection within skipped exons were investigated in the study conducted by Yeo el al. (2004). Human liver was reported to produce the highest number of genes with alternative 3’SS exon AS events at 17%,

followed by the brain at 11% and lastly breast tissue at 4% (Yeo et al., 2004). Additionally, the tissue which produced the highest number of genes with skipped exons was the brain at 26%, followed by the testes at 23% and ovary tissue at 10% (Yeo et al., 2004).

The high proportion of AS events in the liver and brain of humans may be necessary for the differentiation of these tissues. Splicing factors present in these tissues also could be affecting the tissue specific levels of AS seen (Yeo et al., 2004).

Alternative splicing in non-mammalian vertebrates

ESTs or mRNA analysis on whole organism AS rates of non-mammalian vertebrates have been performed and the results are varied (reviewed by Kim et al., 2006). For example, a study by Brett et al. (2002) estimated that across metazoan organisms the rates of AS are similar. In comparison, a study by Kim et al. (2004) suggested that AS rates are higher in mammals compared with invertebrates. The AS rate in chickens is perhaps the most well documented non-mammalian vertebrate species reported (Kim et al., 2004). In chickens the percent of whole organism AS is estimated to be 42%, the highest recorded non-mammalian vertebrate species rate determined thus far (Artamonova et al., 2007; Kim et al., 2007). These results were gathered from 139 chicken ESTs libraries, which included all available tissue types, including embryonic, and cell lines. Kim et al. (2007) reported no cell line or tissue specific AS in chickens (Kim et al., 2007). Using available ESTs, it was determined that out of 162 AS events which can occur in mammals, only 37 (22.9%) were present in frog species (Mudge et al., 2011).

To date, no large scale analysis of tissue specific AS in non-mammalian vertebrates has been conducted. Present studies that have investigated tissue specific AS in these species are gene specific. For example, Venables et al. (2012) showed 4 Fox-regulated exons were spliced in a tissue-specific manner in muscle and heart tissues of *Xenopus laevis*. In chickens, Musclebind-like (MBNL) proteins MBNL1 and MBNL2 were shown to have alternatively spliced isoforms which were highly tissue specific to heat tissue (Terenzi et al., 2010). Tissue specific AS was demonstrated in the lizard (*H. Suspectum*) when two alternatively spliced proglucagon mRNA transcripts were found to only be expressed either in the lizard pancreas or intestine tissues (Chen et al., 1997).

Function of unproductive transcripts

Large scale EST studies show splice variants of various genes produced in different tissues and cell types in many species (McManus et al., 2011; Xu et al., 2002). Some of these SVs do produce functional proteins (reviewed by Nilsen et al., 2010; Lareau et al., 2007). For example, a recent study showed that in human embryonic stem cells (ESC) the AS of the Forkhead box protein P1 gene, *FOXP1*, produced functional proteins (Gabut et al., 2011). FOXP1 is a transcription factor that regulates a number of genes involved in cell proliferation, differentiation and development (Nakae et al., 2003). When the canonical transcript of *FOXP1* is translated, a reduced expression of pluripotency genes and an increase expression of differentiation genes were observed. Alternatively, when exon 18a was included in *FOXP1* mRNA transcript it was termed *FOXP1-ES* and produced a longer protein. FOXP1-ES preferentially binds to specific sets of DNA sequences that are not bound by FOX-P1. Binding of these sequences promotes expression of transcription factors which repress ESC differentiation (Gabut et al., 2011).

FOXP1 and FOXP1-ES are just one example of how AS produces functional proteins. However, many transcripts that are generated fail to do this. In fact, it is proposed that the majority of alternatively spliced mRNAs do not produce functional proteins (Skandalis et al., 2004; Sorek et al., 2004). Many unanswered questions remain, including the purpose of these alternatively spliced transcripts if not to encode for functional proteins. We must ask whether these unproductive transcripts are essentially "non-functional" or do they have a function outside of protein production?

One hypothesis proposes that these alternately spliced mRNAs, although not producing proteins, are functional because they are required during the life cycle of an organism and are activated in a regulatory manner (Sorek et al., 2004). Sorek et al. (2004) proposed this hypothesis based on the investigation of conserved alternatively spliced events between human and mouse genomes. Taken together, a new definition of a functional mRNA transcript has been put forth as one that is required during the life cycle of an organism and is activated in a regulatory manner (Sorek et al., 2004). Additionally, some unproductive transcripts are highly conserved throughout evolution (Iida et al., 2006; Nern et al., 2005). This implies that these transcripts may play a significant role in gene regulation (Nilsen et al., 2010; Skandalis et al., 2010; Artamonova et al., 2007; Lewis et al., 2003). For instance, in humans the gene encoding the ribosomal protein L3 (rpL3), a protein that binds to specific HIV-1 RNA, is alternatively spliced (Bannwarth et al., 2005). The splice variant produced contains an in-frame PTC and is subject to NMD. When canonical rpL3 is overexpressed the level of SVs increases. This increase triggers a negative feedback loop that signals for a decrease of the canonical rpL3. The splice

variant of rpL3 does not encode for a functional protein therefore indicating that AS of this gene regulates its own canonical expression (Cuccurese et al., 2005).

It is difficult to determine the function, if any, in all SVs produced from AS. Even in terms of one specific SV of a gene, we must investigate possible functions independent of protein synthesis. An investigation such as this would have to look at the splice variant's involvement in regulatory pathways at all stages of an organism's development (Skandalis et al., 2010). Predictions on how to determine whether a SV has a function have been developed to address this issue.

The symmetry of an exon is a good start to determine if a variant has function. If an exon is symmetrical, i.e. that is it can be divided by three, the insertion of this exon in a transcript will not disrupt the reading frame. Maintenance of the reading frame may indicate that the transcript could form a functional protein. Alternatively, transcripts that do have a PTC due to AS, may activate the NMD mechanism. The search for functionality of these transcripts can now be narrowed. NMD activation may indicate a negative feedback loop, as seen in the rpL3 gene regulation (Cuccurese et al., 2005).

The evolutionary conservation of a SV is another good indication of possible function (Keren et al., 2010; Skandalis et al., 2010; Lemaire et al, 2004). One study proposes evolutionary conservation in the AS of the fission yeast phosphatidate cytidylyltransferase (CDS1) and its human homolog CHK2 genes. The proteins encoded by these genes play a role in responding to DNA damage at a cell checkpoint. Both CSD1 and CHK2 produce SVs with a frame shift in the open reading frame due to a PTC. It is proposed that these variants act as dominant negatives that play a role in the DNA damage response. Since this variant is found in both yeast and humans in the respective

homologs it has been suggested that this is evidence of an evolutionary conservation AS mechanism (Lemaire et al, 2004).

Another example of evolutionary conservation of AS is observed in the DNA polymerase beta gene, *POLB.* Skandalis et al. (2010) reported that the AS of *POLB* is conserved in primary untransformed fibroblast cell lines of five primate species. The study revealed that one particular SV, exon two skip, was found to be conserved throughout all species investigated (Skandalis et al., 2010). This SV does not produce a functional protein (Simonelli et al., 2009), however, it is still conserved among mammalian species. The AS of *POLB* is the focus of this thesis and will be discussed in detail.

DNA Polymerase Beta (POLB)

DNA Polymerase Beta (POLB) is a highly conserved, constitutively expressed protein involved in DNA repair (Sawaya et al., 1994). Belonging to the Pol X-family species of proteins, POLB is believed to be expressed in the genomes of all vertebrates as a DNA directed DNA polymerase (Uchiyama et al., 2009). POLB is the main enzyme which performs a type of repair called base excision repair (BER), also called gap-filling DNA synthesis (Thompson et al., 2002; Sobol et al., 1996). Specifically, POLB is performs short patch BER, which replaces single nucleotides mistakes but it has been reported to repair gaps up to six nucleotides (Idriss et al., 2001). In order to do this, POLB uses alkaline phosphatase lyase activity to remove the 5’ phosphate of the mismatched nucleotide and will add one nucleotide to the 3’ end of the gap (reviewed in Parsons et al., 2011). When DNA damage is detected, POLB, which is normally located in the cytoplasm, will translocate to the nucleus. POLB is a monomeric 335-amino-acid,

39-kDa polymerase containing two distinct functional domains (Idriss et al., 2002; Abbotts et al., 1988). The first domain, the 8 kDa amino-terminal has DNA binding and lyase activity (Idriss et al., 2002). The second is the 31 kDa carboxyl-terminal domain that functions as the catalytic, also called the polymerase, domain (Simonelli et al., 2009). The catalytic domain contains finger, palm and thumb subdomains which have double stranded DNA binding, nucleotidyl transferase and dNTP selection activities, respectively. These domains and respective subdomains are encoded by four proteolytic domains which correspond to specific exons of the *POLB* gene (Idriss et al., 2002).

Fourteen constitutive exons encode for the single copy *POLB* gene located on chromosome 8p11.2 (NCBI). Exons 1-4 encode for the single stranded binding deoxyribosephosphate (dRP) lyase catalytic activity which allows POLB to process the 5'-phosphate at the end of the gap. Exons 5-7 encode for the double stranded DNA binding domain and exons 8-11 encode a domain that deal with the nucleotidyl transferase activity. Finally, dNTP selection activity of POLB is encoded by exons 12-14 (Idriss et al., 2002). The mRNA sequence conservation of *POLB* is high among vertebrates. A Basic Local Alignment Search Tool (BLAST) shows that compared to human *POLB* mRNA, primate species exhibit >90% nucleotide sequence identity. Other vertebrate species *POLB* mRNA sequence identity percents compared to humans are the domestic dog (92%), the giant panda (92%), the red panda (99%), the African bush elephant (93%), bovine (90%), wild boar (90%), the Chinese hamster (89%), the house mouse (88%), the chicken (80%) and the African clawed frog (76%). All constitutive exons are the same size (number of bps) for the canonical *POLB* mRNA sequences for all

species mentioned here however the genomic DNA shows various size differences in introns.

Alternative splicing and *POLB*

POLB is subject to AS (Bhattacharyya et al., 1999). Studies involving investigation of *POLB* AS patterns have been primarily carried out in humans. However, such splicing has been examined in other mammalian species including primates, cows, sheep, red pandas, dogs, rats and mice, to a lesser extent (Skandalis et al., 2010, Simonelli et al., 2009, Thompson et al., 2002, Skandalis et al., unpublished). A common theme in all these studies has been that AS of *POLB* occurs in high frequency in these species. *POLB* commonly produces alternatively spliced transcripts, also called splice variants or non-canonical transcripts. As seen in humans cell lines the production of *POLB* splice variants were at higher frequencies than the canonical transcript (CT) (CT: 42.6%) (Skandalis et al., 2010). Similar rates were found in normal human bladder tissue and normal bladder cell lines (Thompson et al., 2002). Skandalis et al. (2010) report other primate *POLB* SV frequencies as 52.3 % (*Pongo pygmaeus*), 30.7% (*Gorilla gorilla*), 26.1% (*Pan troglodytes*) and 12.4% (*Macaca fascicularis*).

The types *POLB* SVs have also been investigated (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002). To date, the most common *POLB* SV found in mammals is the loss of exon two (Δ2). *POLB* Δ2 SV is not translated and does not produce a functional protein (Disher et al., 2007, Simonelli et al., 2009). When exon two is lost a premature termination codon (PTC) is created in exon three and results in a truncated 26 amino acid protein. This truncated POLB protein would only contain a portion of the single stranded DNA binding

dRP lyase domain. The function of this splice variant is not yet known although it has been proposed that it may provide a microRNA precursor or act as a dominant negative regulator of normal POLB activity (Skandalis et al., 2010; Thompson et al., 2002). The other common *POLB* splice variants found in humans, in order of prevalence, are loss of exon 11, inclusion of exon α, and loss of exon 4 (Skandalis et al., 2010). Other primates show numerous other SVs in addition to the common ones seen in humans. For example, *Pongo pygmaeus* has the second highest *POLB* SV frequency after humans (Skandalis et al., 2010).

The majority of the SVs recorded in primates were complex multiple exon loss (Skandalis et al., 2010). These SVs consistently had an exon two loss accompanied by a loss of either exon 11, exon 13, exons 12 and 13, exons 3-6, exon 5, exons 3-9, or exon 3-5 and 13. Seven of these splice variants were unique to other non-human primate species, despite the 99% mRNA sequence identity with humans. Complex intron retention and a loss of exon(s) splicing events were also found in some primate species (Skandalis et al., 2010). All data gathered from non-human primate species were collected from untransformed fibroblast cell lines. There are reports that these splicing events are not limited to cells lines and in fact are found in normal human tissues as well (Thompson et al., 2002).

The comparison of SV frequencies and types in cell lines versus tissues is further addressed in Chapter Three. In addition to normal human cell lines and tissues, diseased samples have also been investigated with regard to *POLB* AS.

Alternative splicing of *POLB* and disease

The AS of *POLB* has been linked to disease (Starcevic et al., 2004; Thompson et al., 2002; Bhattacharyya et al., 1999). *POLB* SVs have been identified in various types of human cancers including breast, lung, bladder, gastric, lung, and brain cancer (Simonelli et al., 2009; Thompson et al., 2002; Bhattacharyya et al., 1999; Chyan et al., 1996). The two most common SVs observed have exon two skip or exon 11 skip (Simonelli et al., 2009; reviewed by Sweasy et al., 2005). It has been suggested that a loss of exon 11 results in a loss of important repair functions in cells (Bhattacharyya et al., 1997). For example, Bhattacharyya et al. (1997) showed that HeLa cells transfected with *POLB* Δ11 variants had significantly reduced BER activity. This decrease in BER activity would result in increased amounts of DNA damage and an increase in mutations. It has been suggested that a mutator phenotype could be the result if these mutations occur in genes that normally maintain of genetic stability, like *ras*, p53 and Rb (Bhattacharyya et al., 1997). First proposed by Loeb, the mutator phenotype is an early step in tumour progression (Loeb 2001). This was further supported by a review of small scale studies and it was found that 30% of human tumours express *POLB* SVs (Sweasy et al., 2004).

POLB SVs leading to reduced BER and mutations in cancer related genes are not the entire story. Studies have shown that an over expression of canonical *POLB* transcripts were elevated in breast, colon, prostate cancers, and gastric tumours (Tan et al., 2004; Srivastava et al., 1999). It has been suggested that this increased expression is a result of tumours having higher levels of DNA damage and requiring more functional POLB for repair (Srivastava et al., 1998). Other authors proposed that high levels of POLB destabilize further cellular processes and in turn lead to tumour progression (Sweasy et al., 2003).

Alternative splicing of *POLB* and lifespan

A recent study provided the first demonstration of a correlation between AS and lifespan (Skandalis et al., 2010). This correlation was the first described for any gene and it was found to occur in *POLB*. The authors observed a positive correlation between lifespan and the frequency of unproductive *POLB* SVs. The numbers of types of transcripts, called transcript diversity, also positively correlated with the lifespan of a species. This study was performed in five primate species using untransformed fibroblast cell lines. This is a perplexing and interesting proposal that raises many questions. We must ask, for example, why would an increase in defective POLB, which is used to maintain genomic stability, lead to long lifespan? Additionally, what function do these unproductive transcripts actually have?

A recent review suggests that available evidence shows that reduced BER capacity from decreased levels of POLB plays a role in accelerated rate of aging in Down Syndrome (Patterson et al., 2011). Although this evidence may seem contrary to Skandalis et al. (2010) findings we must first understand the difference between aging and lifespan. Aging is defined as the change in an organism over time and includes changes associated with loss and with gain of function, such as growth and development (Bowen et al., 2004). On the other hand, lifespan is a bit more difficult to define (Gavrilov et al., 1992). The term lifespan can refer to maximal lifespan (the maximal recorded life span of an individual from birth to death, based on empirical evidence), life expectancy (an average number of years an organism is expected to survive), or the longevity of an organism (the average number of years an organism is expected to survive under ideal conditions) (Olshansky et al., 1990). Skandalis et al. (2010) uses the

term “lifespan” when referring to the longevity of an organism. This definition of the term lifespan is also applied and used in this thesis.

There are many genetic factors which can contribute to the lifespan of an organism (Cournil et al., 2001). For example, numerous longevity genetic studies in *Caenorhabditis elegans* have shown that lifespan is linked to cascades of gene activity of insulin-like growth factors (reviewed by Maier et al., 2003). In mammals, it has been shown that tumour protein 53 (p53) can contribute to long lifespan without resulting in cells growing out of control (Maier et al., 2003). Although there are numerous examples of genetic factors which contribute to lifespan of an organism, the idea that a positive correlation exists between lifespan and the production of unproductive transcripts of a DNA repair gene is very intriguing.

It is very important to recognize that Skandalis et al. (2010) were the first to propose this positive correlation between unproductive AS of *POLB* and lifespan. This study leads the way for future investigations. Data was gathered from five primate species cell lines in this study. Many research questions concerning the correlation between lifespan and AS of *POLB* remain. For example, we can ask whether this positive correlation exists in non-mammalian vertebrates, do non-mammalian vertebrates produce SVs of *POLB* at a frequency which correlates positively to the lifespan of the organism and what types of *POLB* transcripts do non-mammalian vertebrates produce?

In addition, Skandalis et al. (2010) used cell lines to gather data. Different types of *POLB* SVs have already been shown to be tissue specific (as discussed above) but it is important to investigate whether tissues produce the same frequency of *POLB* SVs? Is the

POLB AS frequency and positive correlation with lifespan present with regards to various tissue types in mammals?

Lastly, we ask whether the frequencies and diversity of *POLB* AS exist simply due to an organism's overall rate of AS. Based on the large scale AS studies mentioned above, the rates of *POLB* AS could be high in humans as a consequence of higher overall rates of AS. This question can best be addressed in species like humans or other mammals due to the vast collection of data on AS in these species. However, the case is different for non-mammalian vertebrates. Are the frequencies of *POLB* AS in non-mammalian vertebrates due to an overall AS rate of the organism or is *POLB* unique in this context? These three research questions form the basis of the work to be described. The three major aims of the work presented here will address some of these questions.

The first aim of this work is to investigate *POLB* AS in non-mammalian vertebrate tissues. Once the frequencies and diversity of *POLB* transcripts are known in these species, it can be determined whether a positive correlation exists with lifespan, as observed in primate untransformed fibroblast cell lines. The second aim is to compare the data from mammalian cells lines to that of mammalian tissues. The third aim is to address whether the diversity and frequencies of AS of *POLB* is unique to this gene by using the AS information from a second gene for comparison.

METHODS

Tissue collection

All tissue samples were collected following Brock University Animal Care and Use Committee (ACUC) guidelines. Tissue samples from the rabbit, chicken, frog, and bearded dragon were kept in RNA Later (Sigma, #116K0426) and stored at -80 °C. Tissue samples from the gull were collected according to University of British Columbia ACUC. Gull whole organs were extracted, pulverized and stored at -80 °C. Human RNA was purchased from Clonetech (Liver total RNA lot #636531, Brain total RNA lot #1202013).

RNA isolation from tissue samples

Tissues were thawed and placed on a sterile surface. Using a sterile scalpel a small section of the tissue was cut from the whole organ. The remaining tissue sample was immediately placed into RNA later and stored at -80 °C. All RNA isolations from non-mammalian tissues were performed with the RNeasy Mini Kit (Qiagen, #74104) and mammalian tissue RNA isolation with a Norgen RNA isolation kit (#582374). The tissues were homogenized one of three ways:

1. Mortar and Pestle: The tissue sample was bathed in liquid nitrogen. Tissue was gently grinded with a sterile mortar and pestle. After the tissue was sufficiently broken down, 600 μL of lysis buffer was added. Gentle mixing with a pipette created a solution of ground up tissue sample and lysis buffer (hereafter referred to as the “lysate”). The lysate was then used in the next steps of the RNA isolation protocol.

2. Electric Homogenizer: A small piece of tissue was put into a 2 mL flat bottom cryogenic vial and 600 µL of lysis buffer was added; the mixture was then homogenized.
3. Manual Glass Homogenizer: A sterile manual glass homogenizer (KimbleChase, #885303-0002) was placed in an oven for 48 hours at 80 °C prior to use. A small piece of tissue was placed in the tissue grind pestle and 600 µL of lysis buffer was added. The tissue was homogenized and the lysate was then used in the next steps of the RNA isolation protocol.

After homogenization the remaining steps for RNA isolation were followed according to the RNeasy Mini Handbook or Norgen handbook. Both RNA isolation kits are appropriate for fresh, frozen, or RNA Later stabilized animal tissues.

Reverse transcriptase (RT) Polymerase chain reaction (PCR) – First strand cDNA synthesis

cDNA synthesis was carried out using Invitrogen's first strand cDNA Superscript III Synthesis Kit (Invitrogen, #1015963). The procedure was carried out on ice as follows in a final volume of 20 µL. The following was added to a chilled 0.2 mL microcentrifuge tube for 1 reaction: Oligo $dT_{(20)}$ to a final concentration of 2.5 µM, 5 µL of RNA (1-3 ng), deoxynucleotide triphosphates (dNTPs) (0.5 mM), and sterile water up to 13 µL. The mixture heated for 5 minutes at 65 degrees C and immediately placed on ice for 1 min. While remaining on ice the following was added to the mixture: 5x First-Strand Buffer (1X), Dithiothreitol (DTT) (0.1 M), RNA inhibitor (2 units), and Superscript III (10 units). The mixture was incubated for 50 minutes at 50 °C and finally at 70 °C for 15 minutes. The cDNA was stored at -80 °C.

Primers used for initial cDNA PCR and cDNA library

Forward and reverse primers were designed specifically for some of the non-mammalian species used based on the canonical sequences available from the NCBI database. For the non-mammalian species from which the canonical sequences were not known, degenerate primers were designed based on multiple sequence alignment of the most closely related species available. All *POLB* primers were designed to start at exon 1 and end at exon 14. All *TRPV1* primers were designed at the mRNA exon 3 (exon containing start codon) and end at exon 15.

Table 1.1 Primers designed for canonical *POLB* sequences in vertebrate species.

Primer name	Species	Primer sequence (5' → 3')	Sequence length (bps) FxR	Ta
bird-polb-for	Chicken, Gull	ATGAGCAAGCGAAAGGCNCCNCAGGAGA	1008	64.8
bird-polb-rev	Chicken, Gull	CCACTGTATGTANTCAAAGATGTCTTTNTCA		64.8
Emma-polbf	Chicken, Gull, Bearded Dragon, Frog	CAGGAGAGCCTCAACNAGGGNATC	972	68.4
Emma-polbr	Chicken ,Gull, Bearded Dragon, Frog	CTTTGGCTCCCGGTATTTCCACTG		68.4
EmmaDEGPOLBf	Chicken, Gull, Bearded Dragon, Frog	ATGAGCAARMGNAARGCNCCNCARGAGA	1006	69.0
EmmaDEGPOLBr	Chicken, Gull, Bearded Dragon, Frog	ATTCRCTNCGRTCYTTNGGCTCNCGGTA		69.0
Frog-polb-f	Frog	AAGGAAAGCACCACAGGAGAGCCC	997	68.4
Frog-polb-r	Frog	TCATTCGCTGCGARCCTTTGGCTC		68.4
NewtFishPOLBF	Newt, Zebra Fish	GCRCCACARGARWSCCYSAAYGAAGG	993	68.0
NewtFishPOLBR	Newt, Zebra Fish	TCAYTCGCTGCGRAGYAAYCCGTC		68.0
POLBuniF1	Rabbit	CCGCAGGAGACCCTCAACGG	965	63.0
POLBuniR2	Rabbit	CCACTGGATGTAATCAAAAATGTC		63.0
POLBfor1	Human	TGAACACTCTCTGGGGTTCTCG	1121	64.8
POLBrev1	Human	CCAAAGACCCTTACATAGCA		64.8

Table 1.2 Primers designed for canonical *TRPV1* sequences in vertebrate species.

Primer name	Species designed for	Primer sequence (5' → 3')	Sequence length (bps) FxR	Ta
Emma-trpv1f	Chicken, Gull	CAAGGACATGGCTCCAATGGACTC	2077	69.0
Emma-trpv1r	Chicken, Gull	CCACGTGGACCAGTTCACTTCATC		69.0
FrogTRPV1bF	Frog	AAGGCCGTCGCTTTGGGAAGA	2061	68.5
FrogTRPV1bR	Frog	AATCCTCCTTGCCGTCGGGTG		68.5
hTRPV1F	Human	ACTCCTGCCCGACCATCACAGTCA	2048	68.5
hTRPV1R	Human	GACGAGGTGAACTGGACCACCTGG		68.5

Polymerase Chain Reaction (PCR)

PCRs were preformed in final volumes of 20 μL. PCR reactions were performed using two types of Taq Polymerases. Initial PCR from cDNA for *POLB* and *TRPV1* for the frog brain and gull liver were preformed with GoTaq, *TRPV1* from frog liver was preformed with Taq DNA Polymerase with standard 10 X Taq Buffer and all other PCR reactions were preformed with Taq DNA Polymerase with 10X ThermoPol Buffer. All PCR reactions were preformed in a final volume of 20 μL, using Taq Polymerases with their respective buffers, 10 mM Nucleotide solution mix (New England Biolabs), 10 μM forward primer, 10 μM reverse primer, and 2 μL of template.

PCR programs

PCR programs were optimized for each primer pair. The general parameters for all *POLB* primer pairs were 94 °C for two minutes to denature DNA. Next, the reactions underwent 34 cycles of heating to 94 °C for 30 seconds, 56.5-70 °C for 30 seconds then 72-74.8 °C for one minute. A final elongation of 10 minutes at 72 °C was added and the cycling was complete.

PCR programs for all *TRPV1* primers were as follows: 95 °C for two minutes for initial denaturation step followed by 30-35 cycles of heating to 95 °C, 67-70 °C for 25

seconds then 72-74 °C for two minutes. A final elongation time of 5 minutes at 72 °C was completed to finish the cycling and reactions were held at 4°C.

Agarose gel electrophoresis

To visualize PCR products, a 0.7% agarose gel was made with 1X Tris-Acetic acid-Ethylenediaminetetraaceticacid (EDTA) buffer (TAE buffer). Before casting, 20 mg/mL of ethidium bromide was added. 6X loading dye was mixed with the PCR product and loaded into the agarose gel. A DNA ladder was used for a marked of amplicon size. The gel was then visualized under ultraviolet (UV) light and a digital photograph was taken.

Ligation of PCR products

Ligation vector systems used throughout this work were TOPO TA Cloning Kit (Invitrogen, #45-0641) and pGEM-T Easy Vector system (Promega, #326618). Both ligation systems were used for ligation of *POLB* PCR products for the chicken liver. Although both systems produced the same types of transcripts pGEM-T Easy vector system showed a higher efficiency and therefore was used for the remainder of the work.

Ligations made with the TOPO TA Cloning Kit were set up with a final volume of 6 µL. The TOPO Cloning reactions proceeded with 1 µL of dilute salt solution, 1 µL pCR2.1-TOPO, PCR product 2 µL, and 2 µL of water. Reactions were gently mixed and incubated for 5 minutes at room temperature (22-23 °C). Reactions were stored at -20 °C overnight.

pGEM-T Easy vector system ligation protocol was performed according to manufactures specifications. The standard ligation reactions for PCR product required 5 µL 2X Rapid Ligation Buffer, T4 DNA Ligase, pGEM-T Easy Vector, 01-3 µL of PCR

product, T4 DNA Ligase (3 Weiss units), and nuclease-free water was added to a final volume of 10 μL. All ligations were incubated overnight at 4 °C and stored at -20 °C.

Transformation of ligated PCR products into *E.coli*

Ligated PCR products were transformed with competent *E.coli* cells. Two different transformation procedures were used throughout this work, electroporation and heat shock. The transformation procedure determined which of the four types of competent *E. coli* cells was used. The electroporation transformation required electrocompenent *E. coli* cell types JM109 or Electromax DH10B whereas a heat shock transformation used heat shock competent *E. coli* cell types High Efficiency JM109 or DH5α.

Heat shock transformation

Heat shock transformation of competent High Efficiency JM109s or DH5α *E. coli* and ligation product was done using the protocol developed by the De Luca Lab at Brock University (2007). An aliquot of 0.5-2 μL of ligation product was added to 100 μL of competent cells and mixture was incubated on ice for 20 minutes. Cells were heat shocked in a water bath at 42 °C for exactly 45 seconds and then chilled on ice for 5 minutes. One mL of room temperature Super Optimal broth with Catabolite repression (SOC) Medium was added. Cells were cultured for 1 hour at 37°C, shaking at approximately 200 rpm. Cells were then plated on indicator plates.

Electroporation transformation

Electrocompenent JM109 cells were thawed on ice. Fifty μL of cells and 1.25 μL of ligation reaction were added to a cooled electroporation cuvette. The cuvette was incubated on ice for 1 minute and then electroporated with a BioRad Gene Pulser II at 25

μF, 2.5 kV and 200 Ω. One mL of room temperature SOC Medium was added. Cells were cultured for 1 hour at 37°C, shaking at approximately 200 rpm. Electromax DH10B procedure was the same as JM109 cells except that 20 μL of Electromax DH10B cells were used and electroporated at 2.0 kV.

Plating of competent cells

All ligation systems used provided the requirements for blue/white colour screening on indicator plates. Indicator plates were made with LB Agar, Ampicillin, 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal) and Isopropyl-β-D-thio-galactoside (ITPG).

Making indicator plates

LB agar was made by adding 15g/L of Agar (Anachmia, # 02116-300) to LB Media, mixed under heat, and sterilized by autoclaving. The LB agar solution was cooled in a water bath until temperature reached approximately 60 °C. Ampicillin was added to a concentration of 125 μg/mL of LB agar per plate. X-gal and IPTG were added until a concentration of 80 μg/mL and a 1:1 ratio was reached for each reagent respectively, per plate. The plates made in this fashion are hereafter referred to as Ampicillin, X-Gal, IPTG (A/X/I Plates). A/X/I plates were then either used immediately or wrapped in aluminum foil and stored at 4 °C until needed.

Plating amounts and procedure

Aliquots of 50-150 μL of transformed *E.coli* cells were spread over A/X/I LB agar plate surface with a glass spreader. Inoculated A/X/I plates were placed upside-down at 37°C overnight. Resulting white bacterial colonies were then grid plated on A/X/I plates and incubated overnight at 37°C.

Making templates from grid plates

Templates for colony PCR were made with colonies from grid plates. A sterile toothpick was used to pick up a small portion of each colony. This portion was then mixed with 100 μL of autoclaved Milli-Q water and stored in a 0.2 mL microcentrifuge tube. This colony and water mix is hereafter referred to as "templates" and were used in all colony PCRs.

Restriction endonuclease digests of PCR products

All positive colony template PCR products from each tissue sample in all species were digested using restriction endonucleases. All digestions were preformed according to manufacture guidelines.

Table 1.3 Restriction digests of *POLB* PCR products for vertebrate species.

Species	Restriction enzyme	Buffer	Restriction site (5' → 3')	Expected Fragment Sizes Based on Canonical Sequence (bps)
Chicken	**HindIII** (New England BioLabs, #R01048)	Buffer 2 (New England BioLabs, #B7002S)	A|AGCTT	261;384; 383
Gull	**XmnI (PdmI)** (Fermentas, #ER1531)	Buffer Tango (New England BioLabs, #BY5)	GAANN|NNTTC	733;239
Frog	**HindIII** (New England BioLabs, #R01048)	Buffer 2 (New England BioLabs, #B7002S)	A|AGCTT	537;466
Bearded Dragon	**HindIII** (New England BioLabs, #R01048)	Buffer 2 (New England BioLabs, #B7002S)	A|AGCTT	625;381
Rabbit	**EcoRI**	Buffer 3	G|AATTC	428;537
Human	**EcoRI**	Buffer 3	G|AATTC	428;597

Table 1.4 Restriction digests of *TRPV1* PCR products for non-mammalian species.

Species	Restriction enzyme	Buffer	Restriction site (5' → 3')	Expected Fragment Sizes Based on Canonical Sequence (bps)
Chicken	Double digest **XmnI/EcoRI** (Fermentas, #ER1531)/ (New England BioLabs,#1015)	Buffer Tango (New England BioLabs, #BY5)/Buffer 2 (New England BioLabs, #B7002S)	GAANN|NNTTC/ G|AATTC	260; 980;838
			G|AATTC/ C|CATGG	260;406;777;634
	Double digest **EcoRI/NcoI** (New England BioLabs,#1015)/ (New England BioLabs, #R01935)	Buffer 2 (New England BioLabs, #B7002S)/Buffer 3 (New England Biolabs, #B7003S)		
Frog	**TaqαI** (New England BioLabs, #R01498)	Buffer 3 (New England BioLabs, #B7003S)	T|CGA	269;472;493;827
Human	**NcoI** (New England BioLabs, #R01498)	Buffer 3 (New England BioLabs, #B7003S)	C|CATGG	485;653;270;640

SV sequencing and nomenclature

Sequencing of PCR products was carried out by the McGill University and Genome Quebec Innovation Centre. Sequenced transcripts were compared to the CT to determine splicing events.

Splicing events are indicated with the following nomenclature: Skipped exon(s) are indicated by the Δ symbol followed by the name of the skipped exon(s), intron retentions are indicated with the Σ symbol followed by the name of the intron retained, and any partial exon skipping or intron retention are indicated by a "P" meaning "partial", a Δ or Σ symbol, and the number of the partial exon or an "I" to indicated "intron" followed by the partially retained intron number.

Statistics

Statistics were performed using Statistical Analysis Software (SAS). Chi-squared tests were run using the number of splice variants and the number of canonical transcripts to determine whether a significantly statistical difference existed between the AS frequencies of genes within or between tissues. The significance level used for all chi-squared tests were 0.95 (α=0.05)

CHAPTER TWO: Alternative splicing of *POLB* in non-mammals

Goal of Chapter Two

The goal of this chapter is to investigate the alternative splicing (AS) of DNA polymerase beta (*POLB*) and its correlation to lifespan in four non-mammalian vertebrate species. The non-mammalian species in this work include *Gallus gallus domesticus* (chicken), *Larus glaucescens* (the glaucous-winged gull), *Xenopus laevis* (African clawed frog), and *Pogona vitticeps* (bearded dragon). Hereafter, animals will be designated using their common names.

This chapter will address the following questions: What are the splicing patterns of *POLB* in the liver and brain tissues, in these four non-mammalian vertebrate species? Is there tissue specific *POLB* AS patterns? Are the observed *POLB* AS patterns a result of an overall increase in AS in a given tissue type? This last question will be investigated by examining the splicing patterns of a second gene, *TRPV1*, in liver and brain tissues and comparing the results to the AS patterns of *POLB* in these same tissues. Finally, this chapter addresses the question: Does a positive correlation exist between AS of *POLB* and lifespan in the studied species?

This chapter will explore three major hypotheses. Firstly, that *POLB* splicing is not tissue specific in non-mammals. Secondly, that *POLB* and *TRPV1* have similar AS patterns indicating an overall trend in AS in a tissue type. The final hypothesis in this chapter is that *POLB* SV frequency from non-mammalian tissues positively correlate with lifespan.

INTRODUCTION

POLB is conserved in all vertebrate species studied to date (Uchiyama et al., 2009). The AS patterns of *POLB* in mammals are well documented in the literature with data collections compiled largely from human cell lines and several tissues (Skandalis et al., 2010; Simonelli et al., 2009; Skandalis et al., 2004; Thompson et al., 2002). In primate primary untransformed fibroblast cell lines a positive correlation between *POLB* splice variant (SV) frequency and lifespan was observed (Skandalis et al., 2010). Whether this positive correlation exists in non-mammalian vertebrates has yet to be examined. If this correlation is found in non-mammalian vertebrates, it may suggest that the production of unproductive *POLB* SVs is an evolutionarily conserved mechanism in vertebrates. The research work presented here shows the AS patterns of *POLB* from three major non-mammalian vertebrate groups: avian, amphibian and reptilian.

Two types of avian species were investigated: one domesticated (chicken) and one wild (gull). This was done to observe any influence of domestication on the AS patterns of *POLB*. Chickens have been subject to strong human driven domestication and have evolved with genetic adaptations (Rubin et al., 2010). A recent study reported that *Gallus gallus domesticus* have more than 7,000,000 single nucleotide polymorphisms, almost 1,300 deletions, and some differences in the thyroid stimulating hormone receptor (TSHR) compared to a major wild ancestor (Rubin et al., 2010). It is possible that domestication has influenced AS in the chicken. Should this be the case, then AS data from the chicken may not be a reliable indicator of AS patterns in the avian group.

To investigate AS of *POLB* in amphibians, liver and brain tissue was obtained from *Xenopus laevis* (frog). Tissues from an amphibian second species, *Notopthalmus*

viridescens (newt), were also tested but cloning of *POLB* transcripts was unsuccessful. Finally, tissues from the bearded dragon were obtained to investigate AS of *POLB* in reptilians. The lifespan expectancy of each of these species ranges from 12-30 years (Table 2.1).

Table 2.1 Non-mammalian species lifespan data.

Species	Lifespan (years)	Reference
Chicken	30	AnAge database
Gull	32	AnAge database
Frog	15	Nuno *et al*., 2010
Bearded Dragon	12	Reusch *et al*., 2009

To determine whether the *POLB* AS patterns observed are the result of a tissue specific rate of AS, a second gene, described below, was introduced for comparison. The comparison gene had to share common features with *POLB* while simultaneously possessing stark differences. It was highly conserved across species, similar to *POLB*. This conservation had to be present not only in nucleotide sequence homology but also in the function of the canonical protein. The comparison gene also had a wide tissue distribution, like *POLB*. This thesis investigated tissue specific AS patterns and therefore, the second gene had to be expressed in all tissue types tested. Finally, it is imperative that, like *POLB*, the comparison gene must undergo AS. The most important difference between the comparison gene and *POLB* was that it was not a DNA repair gene. This difference was essential because AS patterns of *POLB* might be common to DNA repair genes.

Based on these requirements, the gene used as a comparison for the AS of *POLB* was the transient receptor potential cation channel subfamily V member 1 (*TRPV1*) gene. *TRPV1* is highly conserved across species, has a wide tissue distribution, undergoes AS

and is not a DNA repair gene (Cavanaugh et al., 2011, Pingle et al., 2007). Tissue specific AS frequencies and diversity of *TRPV1* were investigated in two of the four non-mammalian species: the chicken and the frog. Although amplification and cloning of *TRPV1* was attempted in all non-mammals, time constraints permitted that only *TRPV1* transcripts from the chicken and frog were examined.

TRPV1

The TRP superfamily of proteins is a group of cation channels (Moiseenkova-Bell et al., 2008). The TRPV1 protein was the first family member to be identified from the superfamily of vanilloid type of TRP nonselective cation channels (Caterina et al., 1997). TRPV1 is also called the capsaicin receptor because of its well studied sensitivity and activation to capsaicin, the natural product of the hot capsaicin peppers. TRPV1 is sensitive to heat (>43°C), protons (pH<6.3), vanilloids, acids, and endogenous proinflammatory substances (Caterina et al., 1997; Caterina et al., 2000). Similar to the other members of the TRP superfamily, TRPV1 is cation nonselective having nearly equal selectively for Na^+, K^+, Li^+, Cs^+, and Rb^+ ions (Caterina et al., 1997).

In humans, TRPV1 is an 838- amino acid, 95-kDa protein. Similar to the other members of the TRP family, TRPV1 has six transmembrane (TM) domains. Between the fifth and the sixth TM domain is a pore-forming region which forms the ion pore to allow selectively of ions (Tominaga et al., 2000). The first 400 amino acids encode for the carboxyl terminus of the protein which contains three ankyrin-repeat domains which mediate protein-protein interactions (Michaely et al., 2002). Amino acids 434-680 encode for the six TM domains: two of which are extra cellular domains, two are

cytoplasmatic domains and one is a pore forming domain. Functional TRPV1 proteins form as tetramers existing as homo- or heteromultimers.

The TRPV1 gene is highly conserved across species with diverse tissue distribution (Cavanaugh *et al*., 2011, Pingle et al., 2007). The highest expression levels have been reported in neuronal tissues including the dorsal root ganglia and trigenminal ganglia in mammals (Pingle et al., 2007). Additional TRPV1 expression has been detected in nonneuronal tissues such as the bladder, liver and macrophages (Tominaga et al., 2005).

The mRNA sequence for the canonical *TRPV1* gene, which codes for the TRPV1 protein, has been described in humans (Caterina et al., 1997), rat (Suzuki et al., 1999), mouse (Caterina et al., 1999), chicken (Jordt et al., 2002), frog (Bohlen et al., 2010), and several other species (NCBI). The human *TRPV1* mRNA is 4074 base pairs. In all species described there are 17 exons in the mRNA transcript and 15 in the coding sequence (CDS). *TRPV1* undergoes AS to form several SV forms (Gracheva et al., 2011; Tian et al., 2006; Wang et al., 2004; Schumacher et al., 2000).

To date there have been five SVs of the *TRPV1* mRNA identified. The first SV, VR.5'sv, was observed in rat tissues (Schumacher et al., 2000). This SV lacks exons 1-5 and 7. Starting translation at exon 6, the protein produced lacks the N cytoplasmic domain and the first two ankyrin repeats. It is not activated by capsaicin, protons or temperature and does not produce and a functional channel (Schumacher et al., 2000). It has been proposed that VR.5'sv protein functions as a negative modulator to canonical TRPV1 activation (Eilers et al., 2007).

Wang et al. (2004) described the second *TRPV1* SV, TRPV1β in dorsal root ganglia tissue of mice. This variant is missing the last 30 bps in exon seven resulting in a deletion of 10 amino acids between the third ankyrin repeat and the first TM domain. When expressed alone TRPV1β is not functional. When this variant is co-expressed with canonical TRPV1, TRPV1β inhibits expression of TRPV1 in mice, thereby acting as a dominant negative regulator (Wang et al., 2004).

Lu et al. (2005) described a *TRPV1* variant with an exon seven deletion. This variant was found to be expressed in human tissues and is called hTRPV1b (Lu et al., 2005). Interestingly, this variant produces a functional protein despite missing 60 amino acids from the cytoplasmic domain, the first two ankyrin repeats and part of the third one. hTRPV1b has been shown to be activated by temperature but not capsaicin or protons (Lu et al., 2005).

TRPV1(VAR) is the name of the fourth known *TRPV1* SV described by Tian et al. (2006). It is the first of two described *TRPV1* variants with an intron retention AS event. This variant retains 101 bps between exons five and six that contains a PTC at amino acid position 254. The PTC results in a truncated protein of 253 amino acids. This variant has been shown to act as a dominant negative regulator or potentiator of canonical TRPV1 expression, depending on cell type used (Tian et al., 2006).

Finally Gracheva et al. (2011) revealed the newest addition of *TRPV1* SVs, TRPV1-S, was recently found in vampire bats. TRPV1-S is the second *TRPV1* variant described with an intron retention AS event. This 23 bp inclusion after exon 14 contains a stop codon which leads to the production of a short (-S) transcript. Production of the TRPV1-S protein results in a channel that has a lower thermal activation threshold (~30 °

C). It has been proposed that this channel allows vampire bats to physiologically tune their thermosensory nerve fibers (Gracheva et al., 2011).

Transcripts of *POLB* and *TRPV1* were collected and characterized from liver and brain tissues from all four non-mammalian species. In all cases, except the brain tissue from the bearded dragon, *POLB* showed SV frequencies of $\leq$3.2%. AS of *POLB* was shown to be tissue specific in the gull and the bearded dragon. No statistical difference existed in the AS of *POLB* in the brain tissues between the chicken and the gull, whereas a statistically significant difference was observed in liver tissues. *TRPV1* and *POLB* splicing demonstrated statistically significant differences in both the chicken and the frog tissues.

RESULTS

AS of *POLB* in chicken tissues

POLB cloned transcripts were successfully amplified from chicken liver and brain tissues (Figure 2.1, lanes 4 and 6). The resulting size of the PCR product was approximately 1008 bps. (Table 1.1). A total of 155 and 74 *POLB* transcripts were characterized from the chicken liver and brain tissues, respectively.

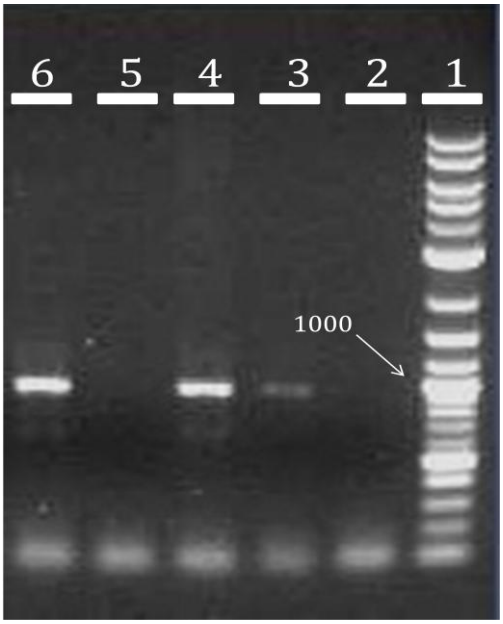

Figure 2.1 *POLB* PCR products amplified from chicken liver and brain tissue cDNA. Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lane 3 contains a positive control, in which mammalian cDNA was used with mammalian *POLB* primers. Lanes 4 and 6 show *POLB* PCR product from the chicken liver and brain tissues, respectively.

AS of *POLB* in chicken liver tissue

Of the 155 *POLB* transcripts characterized from chicken liver tissue five were SVs. The *POLB* SV frequency in chicken liver was found to be 3.2% (Table 2.4). The *POLB* transcript diversity was 3 (Table 2.3). A *POLB* transcript with intron one retention was characterized, ΣI1 (Table 2.2, Figure 2.4). The inclusion of intron one disrupts the reading frame and creates a PTC in exon two. Of the 155 *POLB* transcripts from the chicken liver this SV was detected three times (frequency of 1.9%) (Figure 2.2). The only other *POLB* SV characterized in the chicken liver tissue was a partial 48 bp skip of exon 13, $\Delta p13_{48}$ (Figure 2.3). This partial exon skip does not disrupt the reading frame. Two of these SVs were found out of the 155 total characterized (frequency of 1.3%) (Figure 2.2).

AS of *POLB* in chicken brain tissue

Of the 74 *POLB* transcripts from chicken brain two were SVs of one type, ΣI1 (Figure 2.3). The SV frequency and transcript diversity in the brain tissue were 2.8% and two, respectively (Table 2.3). No statistically significant difference was observed in *POLB* SV frequency between the two tissue types (x^2=0.03, d.f=1, p=0.85, α=0.05) (Table 2.3).

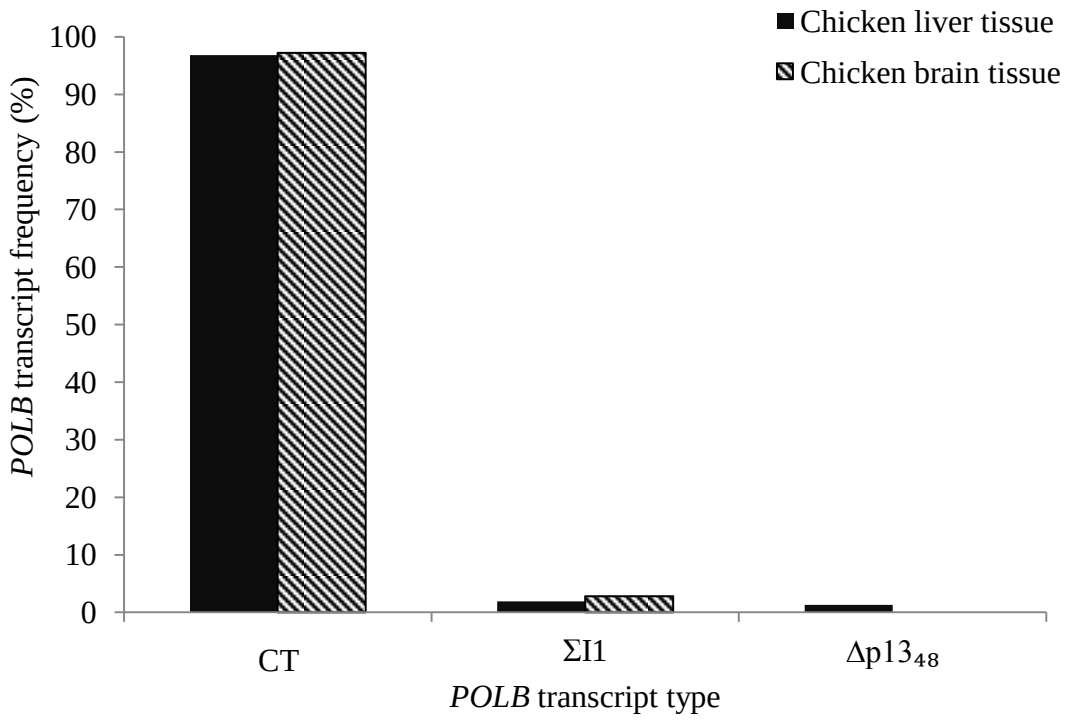

Figure 2.2 *POLB* CT and SV frequencies of the chicken liver and brain tissues.

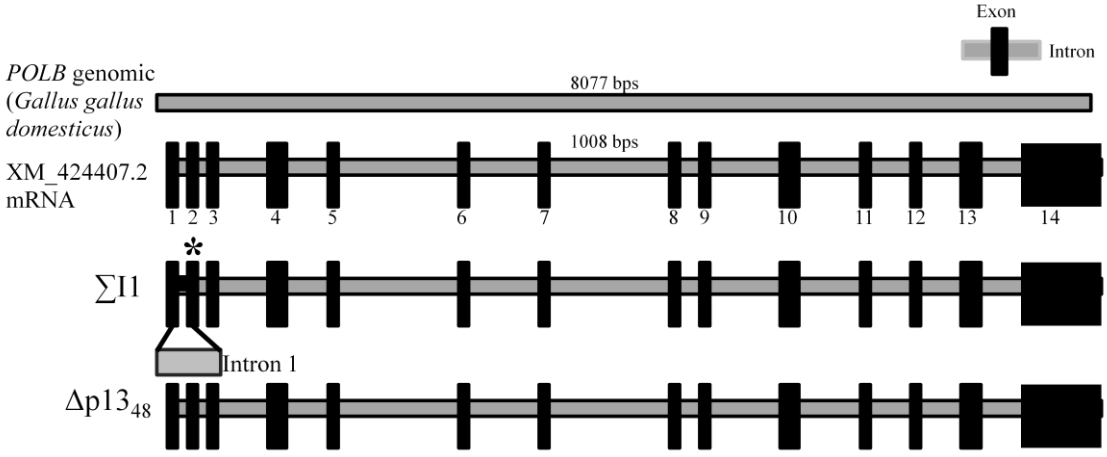

Figure 2.3 *POLB* transcript types characterized from the chicken liver and brain tissues. * Indicates location of a PTC.

Table 2.2 *POLB* ΣI1 from the liver and brain tissues in the chicken.

Name	Sequence
Intron 1	GGTATGGGGGGGCGCGGGGAGGGGGCGTCCTGGGATTCGGCTCGG CTGTTCCAACTGTGTTTCTCCCCGTA

Table 2.3 *POLB* frequency of AS and transcript diversity in chicken liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	150	3.2	3
	Splice variant	5		
Brain	Canonical transcript	70	2.8	2
	Splice variant	2		

Splice variant frequency: the relative frequency of splice variants compared to the total number of transcripts characterized

Transcript diversity: The number of different types of transcripts observed (including CT)

AS of *TRPV1* in chicken tissues

TRPV1 cloned transcripts were successfully amplified from the chicken liver and brain (Figure 2.4, lanes 6 and 8). The resulting size of the PCR product was approximately 2077 bps (Table 1.2). A total of 91 and 103 *TRPV1* transcripts were characterized from the chicken liver and brain tissues, respectively.

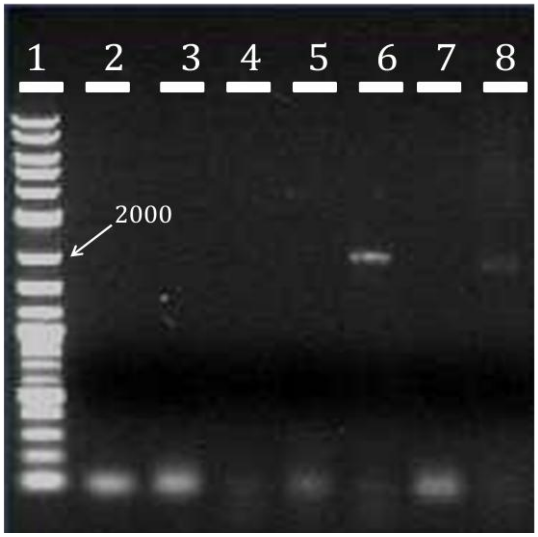

Figure 2.4 *TRPV1* PCR products amplified from chicken liver and brain tissue cDNA. Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lanes 3-5 show lack of amplification of *TRPV1* from cDNA from the bearded dragon. Lane 6 shows successful amplification of *TRPV1* from cDNA from the chicken brain. Lane 7 shows lack of amplification of *TRPV1* from cDNA from the chicken muscle. Lane 8 shows successful amplification of *TRPV1* from cDNA from the chicken liver.

AS of *TRPV1* in chicken liver tissue

Of the 91 *TRPV1* transcripts characterized in the chicken liver tissue 71 were SVs. The SV frequency and transcript diversity of *TRPV1* was 78% and seven, respectively (Table 2.4). The most common SV observed in this tissue type was Δ12 (Figure 2.6). This transcript accounted for 56 of the total 91 transcripts (frequency of 62%) and is 1911 bps in length (Figure 2.5). Exon 12 of *TRPV1* is 166 bps and the skip of this exon resulted in a PTC. The second most common *TRPV1* SV in the chicken liver was $\Delta p4_{63}$ (Figure 2.6). The $\Delta p4_{63}$ SV would result in a 2015 bps transcript. This SV occurred at a frequency of 5.5% (5/91) (Figure 2.5). The skip of the partial exon four did not result in a PTC. Another skip of a partial exon was observed in the *TRPV1* SV, $\Delta p8_{50}$ (Figure 2.6). This splicing event would result in a total *TRPV1* transcript size of 2027 bps. This SV occurred at a frequency of 3.3% (3/91) (Figure 2.5). The skip of the partial exon eight resulted in a PTC in the $\Delta p8_{50}$ *TRPV1* transcript.

A second complete exon skip event was observed in the *TRPV1* SV Δ6 (Figure 2.6). This SV occurred at a frequency of 1% (1/91) (Figure 2.5). Exon six of the canonical chicken *TRPV1* transcript is 141 bps in length and the skip of this exon did not result in a PTC. Another type of *TRPV1* splicing event observed was multiple exon skipping, specifically exons five and six (Figure 2.6). This Δ5,6 *TRPV1* SV occurred at a frequency of 2.2% (2/91) (Figure 2.5). No PTC was created from the multiple skip of exons five and six.

Finally, one intron retention splicing event of intron 15 (∑I15) was observed in one of the 91 transcripts (frequency of 1.0%) (Figure 2.5 and 2.6). This intron retention would result in a total transcript length of 2607 bps and the creation of a PTC.

AS of *TRPV1* in chicken brain tissue

Of the 103 *TRPV1* transcripts characterized in the chicken liver six were SVs. The SV frequency and transcript diversity of *TRPV1* in the chicken liver was 5.8% and four, respectively (Table 2.4). The types of SVs observed were $\Delta p4_{63}$, Δ6 and Δ12, all of which occurred at a frequency of 1.9% each (2/103) (Figure 2.5 and 2.6). A statistically significant difference was found between chicken liver and brain tissues in *TRPV1* SV frequencies (Table 2.4) (x^2=107.7, d.f=1, p=<0.0001, α=0.05)

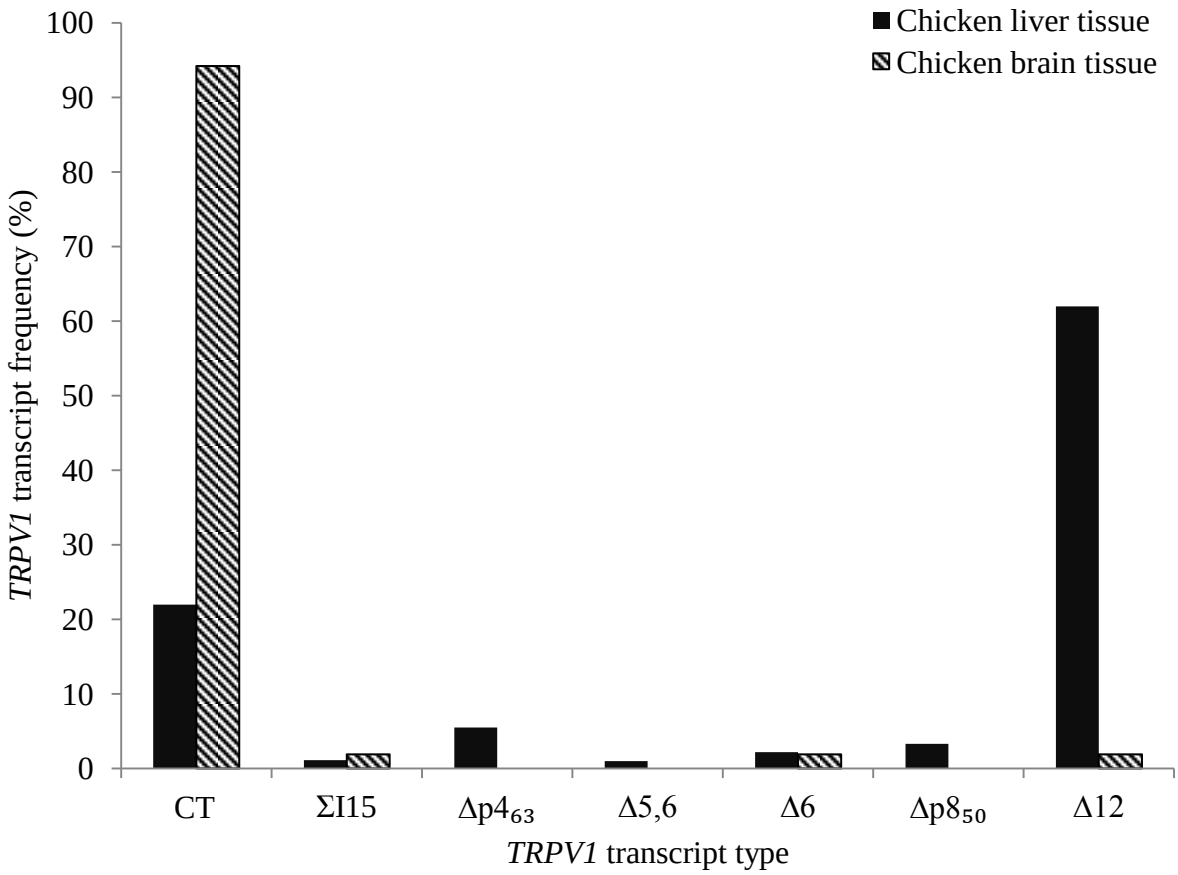

Figure 2.5 *TRPV1* CT and SV frequencies of the liver and brain tissues in the chicken.

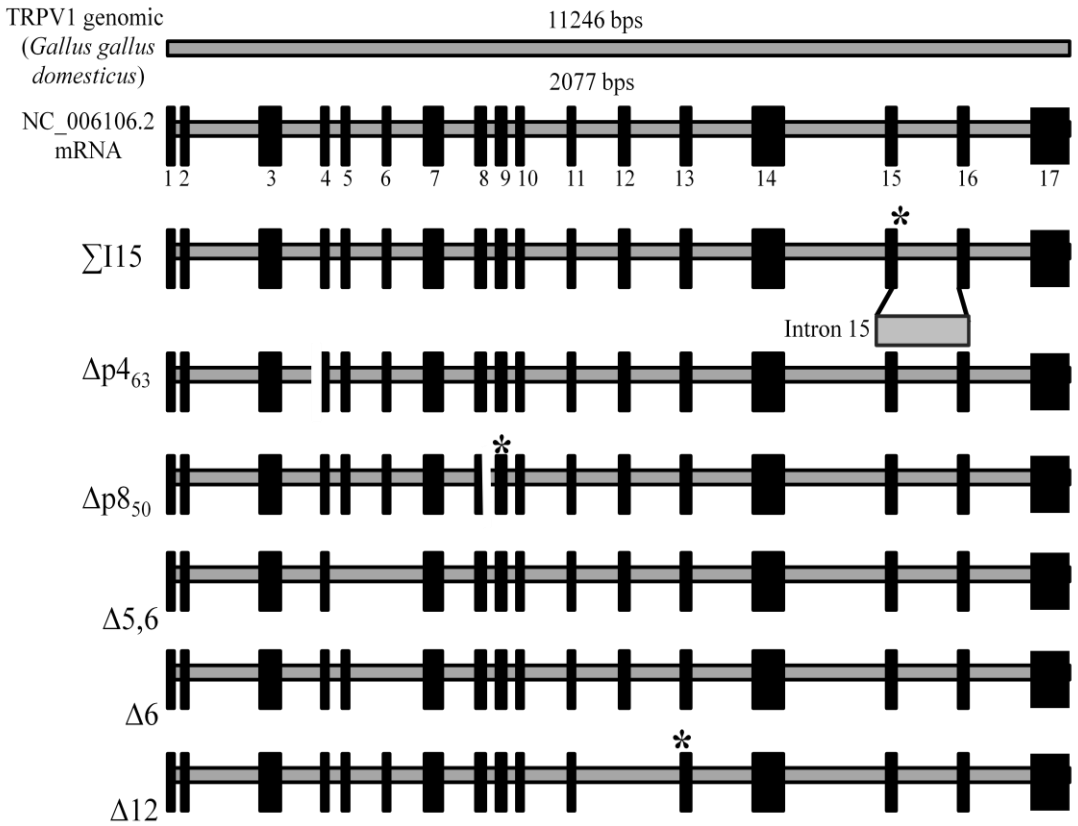

Figure 2.6 *TRPV1* transcript types characterized from the chicken liver and brain tissues. * Indicates location of a PTC.

Table 2.4 *TRPV1* frequency of AS and transcript diversity in chicken liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity (#) in tissue type
Liver	Canonical transcript	20	78.0	7
	Splice variant	71		
Brain	Canonical transcript	100	5.6	4
	Splice variant	6		

***POLB* and *TRPV1* SV frequencies in the liver and brain tissues of the chicken**

There was no statistically significant difference found between *POLB* SV frequency in chicken liver compared to brain (x^2=0.03, d.f=1, p=0.85, α=0.05) (Figure 2.7). *TRPV1* does show statistically significant difference between tissues types in both SV frequency and transcript diversity (x^2=107.7, d.f=1, p=<0.0001.,α=0.05) (Figure 2.7). Statistically significant differences were found between *POLB* and *TRPV1* SV frequencies in the liver and the brain tissue, respectively (x^2=0.8300, d.f=1, p=0.3623, α=0.05) (x^2=150.2, d.f=1, p=<0.0001, α=0.05) (Figure 2.7)

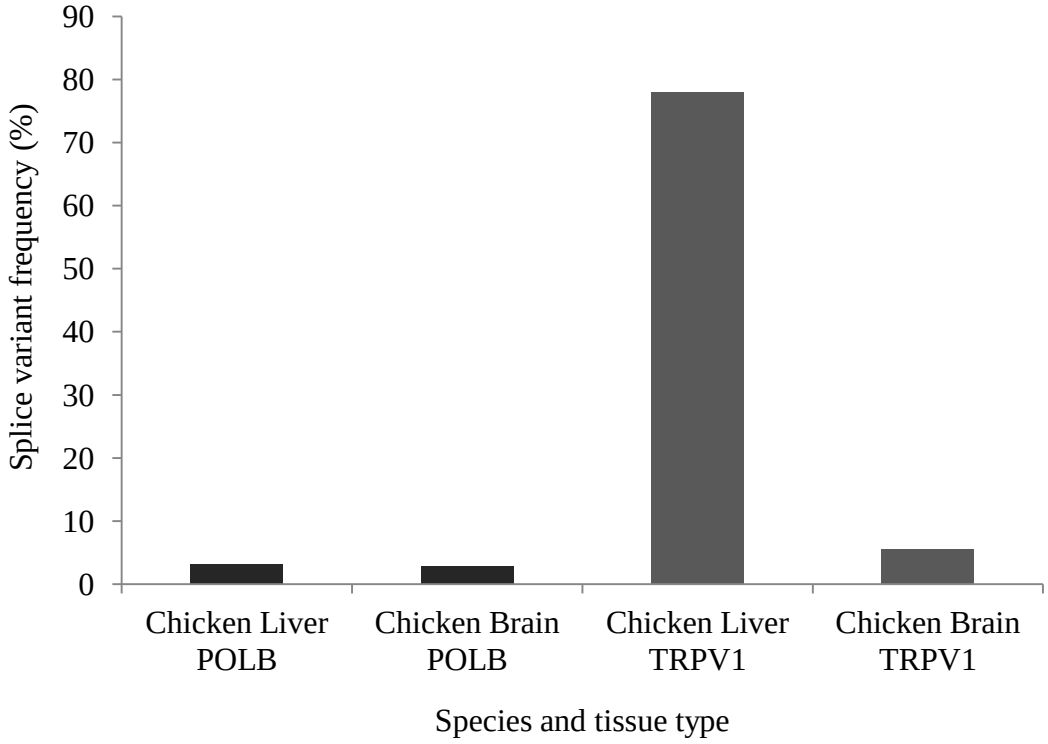

Figure 2.7 SV frequencies of *POLB* and *TRPV1* in the chicken liver and brain tissues.

AS of *POLB* in gull liver tissues

POLB cloned transcripts were successfully amplified from gull liver and brain tissues (Figure 2.8, lane 2 and 2.9, lane 8). The resulting size of the PCR product was approximately 972 bps (Table 1.1). A total of 98 and 73 *POLB* transcripts were characterized from the gull liver and brain tissues, respectively.

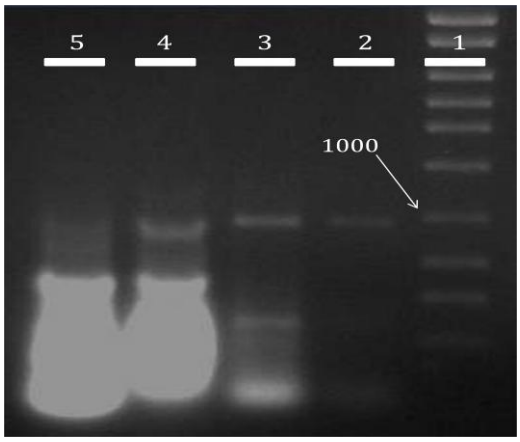

Figure 2.8 *POLB* PCR products amplified from gull liver tissue cDNA. Lane 1 shows the DNA size marker. Lane 2 shows successful amplification of *POLB* using emmaPOLBf x emmaPOLBr. Lane 3 shows successful amplification of *POLB* using emmaPOLBf xPOLBDEGR. Lane 4 shows successful amplification of *POLB* using POLBDEGF x emmaPOLBr. Lane 5 shows successful amplification of *POLB* using POLBDEGF x POLBDEGR. The *POLB* transcript PCR product from lane 2 was used for all other gull liver *POLB* experiments.

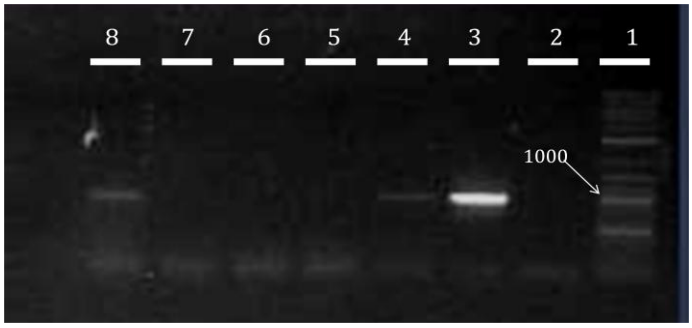

Figure 2.9 *POLB* PCR products amplified from gull brain tissue cDNA. Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lane 3 contains a positive control, in which gull liver colony template was used. Lane 4 shows successful amplification of *POLB* from cDNA from the chicken brain. Lanes 5, 6 and 7 shows lack of amplification of *POLB* from cDNA from the gull liver, scaap brain and mall liver, respectively. Lane 8 shows successful amplification of *POLB* from cDNA from the gull brain.

AS of *POLB* in gull liver tissue

Of the 98 *POLB* transcripts characterized from the gull liver tissue one was determined to be a SV. This *POLB* SV was a partial intron inclusion of 83 bps of intron seven ($\Sigma pI7_{83}$) and occurred at a frequency of 1.0% (1/98) (Figure 2.10 and 2.11). This splicing event resulted in a PTC in exon four of the *POLB* transcript. A partial 972 bps of the canonical gull *POLB* transcript was cloned and characterized from this data (NCBI#: JN620209.1).

AS of *POLB* in gull brain tissue

Of the 72 *POLB* transcripts characterized from the gull brain two were SVs of two types. The *POLB* SV frequency in the gull brain was 2.8% with a transcript diversity of three (Table 2.5). The first type of SV characterized was a complex splicing event involving exons two and three skipping with an inclusion of 83 bps of intron 7 ($\Sigma pI7_{83}\Delta 2,3$) (Figure 2.11). This SV was observed in one out of 72 transcripts (frequency of 1.4%) (Figure 2.10). The skip of *POLB* exons two and three resulted in a PTC in exon four. The second *POLB* SV observed had a frequency of 1.4% (1/72) and was characterized by an inclusion of 83 bps of intron seven (Figures 2.10 and 2.11). The $\Sigma pI7_{83}$ SV contains a PTC as a result of the partial intron retention. There a statistically significant difference in *POLB* SV frequency between gull liver and brain tissues (x^2: 0.7394, d.f=1, p=0.3898, α=0.05) (Table 2.5).

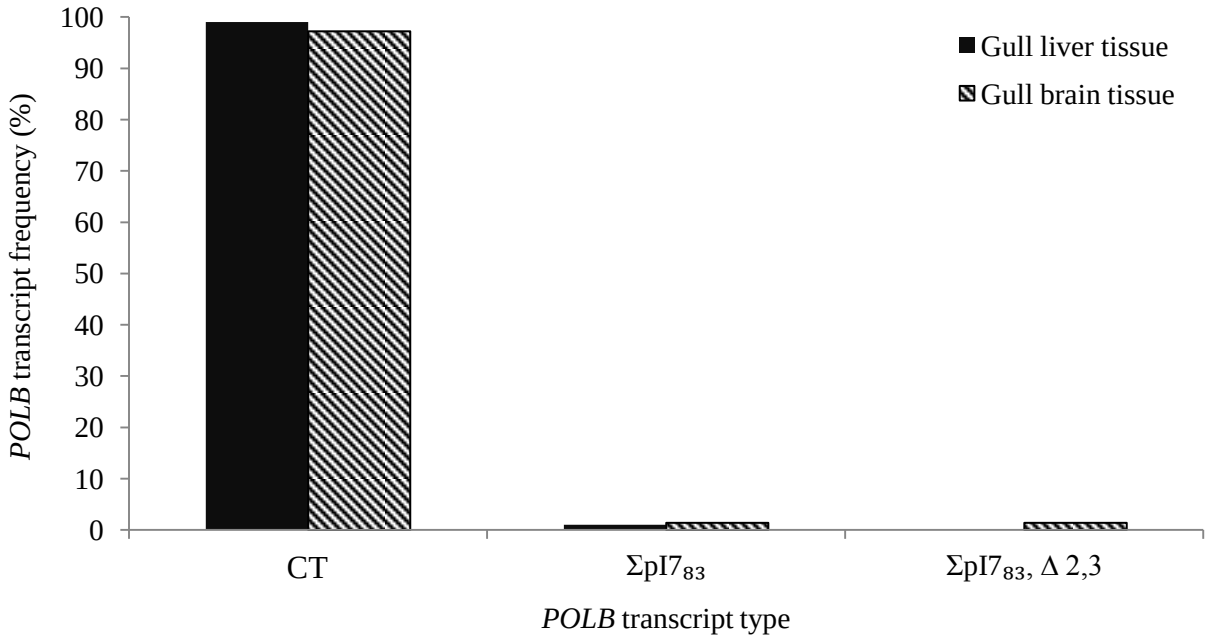

Figure 2.10 *POLB* CT and SV frequencies of the liver and brain tissues in the gull.

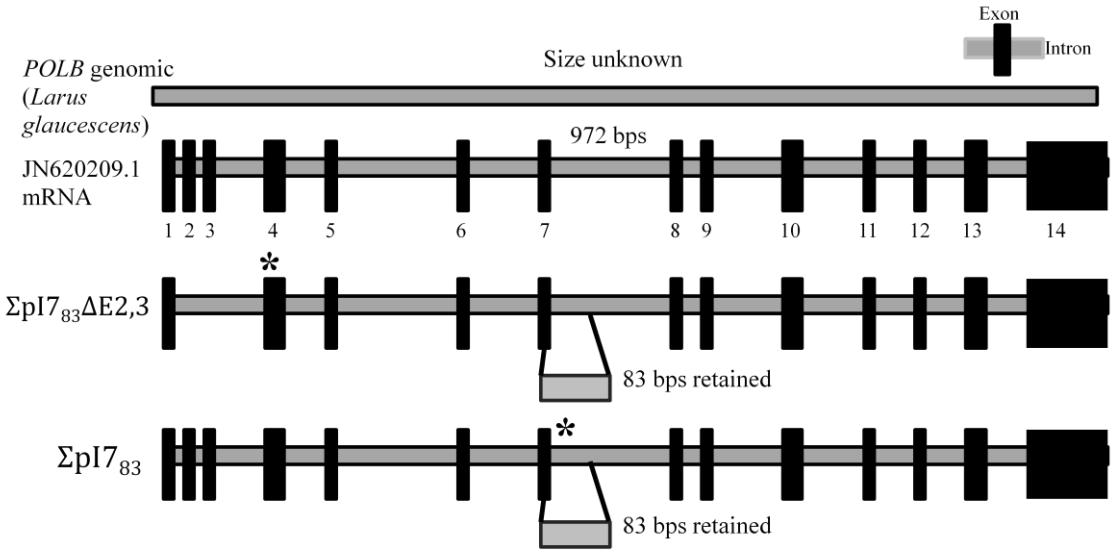

Figure 2.11 *POLB* transcript types characterized from the gull liver and brain tissues.
* Indicates location of a PTC.

Table 2.5 *POLB* frequency of AS and transcript diversity in gull liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	97	1	2
	Splice variant	1		
Brain	Canonical transcript	70	2.8	3
	Splice variant	2		

AS of *POLB* between avian species liver and brain tissues

No statistically significant difference in *POLB* SV frequency was observed between brain tissues of avian species (x^2: 0, d.f=1, p=1.0, α=0.05) (Figure 2.12). In the liver tissues of the avian species a statistically significant difference was observed (x^2: 1.2613, d.f=1, p=0.2614, α=0.05) (Figure 2.12).

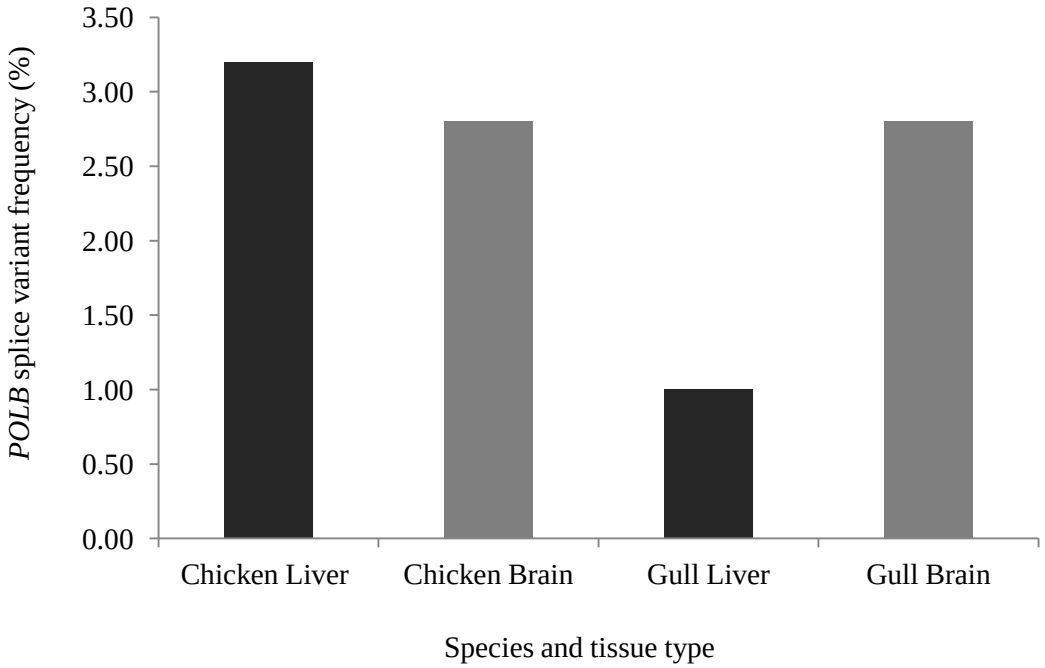

Figure 2.12 *POLB* SV frequencies of the liver and brain tissues of avian species. In all tissue types the *POLB* SV frequency was ≥3.2% in avian species.

AS of *POLB* in frog tissues

POLB cloned transcripts were successfully amplified from frog liver and brain (Figure 2.13 lane 3 and 2.14 lane 5). The resulting size of the PCR product was approximately 1003 bps (Table 1.1). A total of 110 and 127 *POLB* transcripts were characterized from the frog liver and brain, respectively.

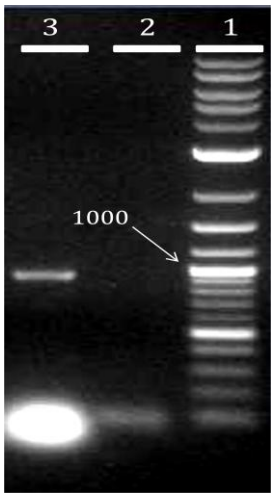

Figure 2.13 *POLB* PCR products amplified from frog liver tissue cDNA Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lane 3 shows successful amplification of *POLB* from cDNA from the frog liver.

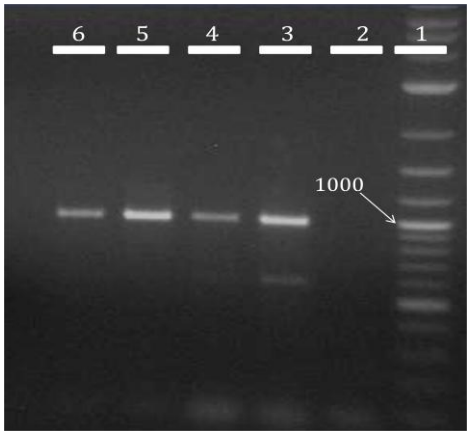

Figure 2.14 *POLB* PCR products amplified from frog brain tissue cDNA. Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lanes 3 and 4 show successful amplification of *POLB* from bearded dragon liver cDNA. Lanes 5 and 6 show successful amplification of *POLB* from the frog brain cDNA.

AS of *POLB* in frog liver tissue

Of the 110 *POLB* transcripts characterized the frog liver one SV was observed. The *POLB* SV frequency in the frog liver was 0.9% with a transcript diversity of two (Table 2.6). The single type of *POLB* SV observed was retention of 681 bp of intron two ($\sum pI2_{681}$) (Figure 2.16). This SV occurred at a frequency of 0.9% (1/110) (Figure 2.15). The retention of this partial intronic sequence created a PTC in the transcript.

AS of *POLB* in frog brain tissue

Of the 127 *POLB* transcripts characterized the frog brain two SVs were observed. The *POLB* SV frequency in the frog brain was 1.6% with a transcript diversity of three (Table 2.6). $\sum pI2_{616}$ was one of the SV types observed, occurring at a frequency of 0.8% (1/127) (Figure 2.15 and 2.16). The second type was a multiple exon skip splicing event of exons two and three (Δ2,3) (Figure 2.16) . The Δ2,3 *POLB* SV occurred at a frequency of 0.8% (1/127) (Figure 2.15). Exons two and three skipping resulted in a 125 bp shorter transcript compared to the canonical transcript and the creation of a PTC. No statistically significant difference was found in *POLB* SV frequencies between frog liver and brain tissues (x^2=0.21, d.f=1, p=0.65, α=0.05).

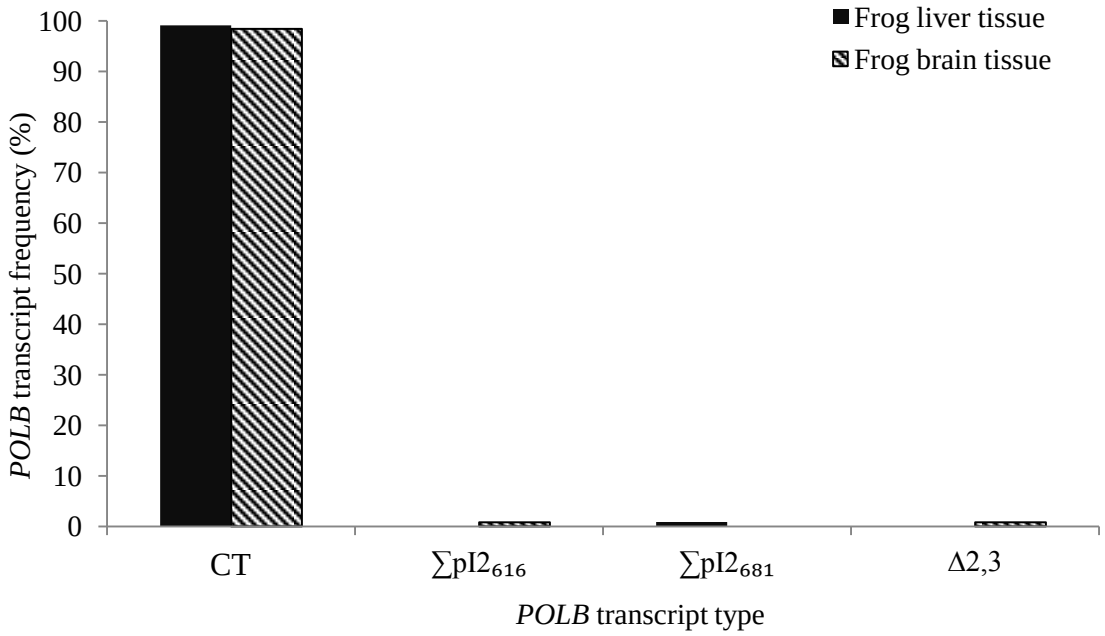

Figure 2.15 *POLB* CT and SV frequencies of the liver and brain tissues in the frog.

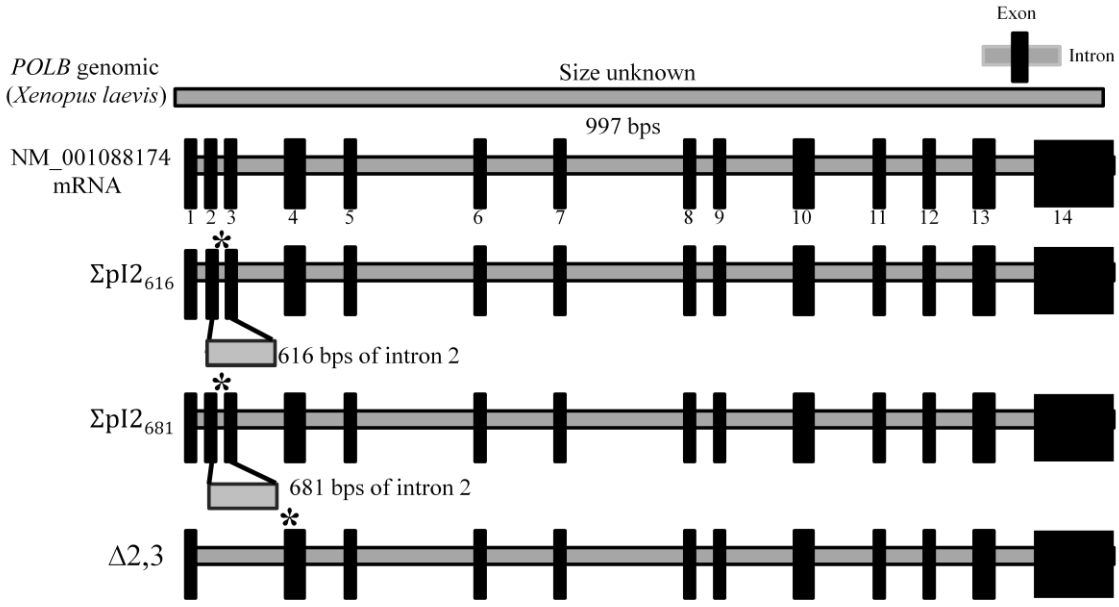

Figure 2.16 *POLB* transcript types characterized from the liver and brain tissues of the frog. * Indicates location of a PTC.

Table 2.6 *POLB* frequency of AS and transcript diversity in frog liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	109	0.9	2
	Splice variant	1		
Brain	Canonical transcript	125	1.6	3
	Splice variant	2		

AS of *TRPV1* in frog tissues

TRPV1 cloned transcripts were successfully amplified from frog liver and brain (Figure. 2.17, lane 3 and 2.18, lane 5). The resulting size of the PCR product was approximately 2061 bps (Table 1.2). A total of 112 and 95 *TRPV1* transcripts were characterized from the frog liver and brain, respectively.

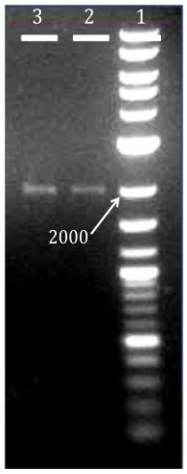

Figure 2.17 *TRPV1* PCR products amplified from frog liver tissue. Lane 1 shows the DNA size marker. Lanes 2 and 3 are duplicate PCR reactions which show successful amplification of *TRPV1* from the frog liver cDNA.

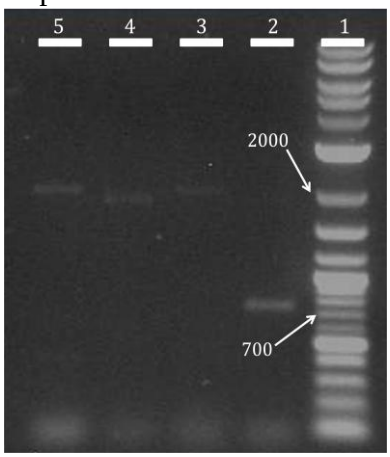

Figure 2.18 *TRPV1* PCR products amplified from frog brain tissue. Lane 1 shows the DNA size marker. Lane 2 shows lack of amplification of *TRPV1* using TRPV1for x TRPV1rev. Lane 3 shows successful amplification of *TRPV1* using TRPV1for x TRPV1bREV. Lane 4 shows successful amplification of *TRPV1* using TRPV1bFOR x TRPV1rev. Lane 5 shows successful amplification of *TRPV1* using TRPV1bFOR x TRPV1REV.

AS of *TRPV1* in frog liver tissue

Of the 112 *TRPV1* transcripts characterized from the frog liver none were SVs. The *TRPV1* SV frequency and transcript diversity of *TRPV1* in the frog liver was 0% and one, respectively.

AS of *TRPV1* in frog brain tissue

Of the 95 *TRPV1* transcripts characterized from the frog brain four were SVs. The *TPRV1* SV frequency and transcript diversity in the frog brain was 4.1% and two, respectively (Table 2.7). The single type of SV observed was an exon skip event resulted in exon 13 skip (Δ13) (Figure 2.20). This *TRPV1* SV occurred at a frequency of 4.2% (4/95) (Figure 2.19). Exon 13 skip in the *TRPV1* transcript resulted in a PTC in exon 14. There is a statistically significant difference in *TRPV1* SV frequencies between frog liver and brain tissues (x^2: 4.3719, d.f=1, p=0.0365, α=0.05).

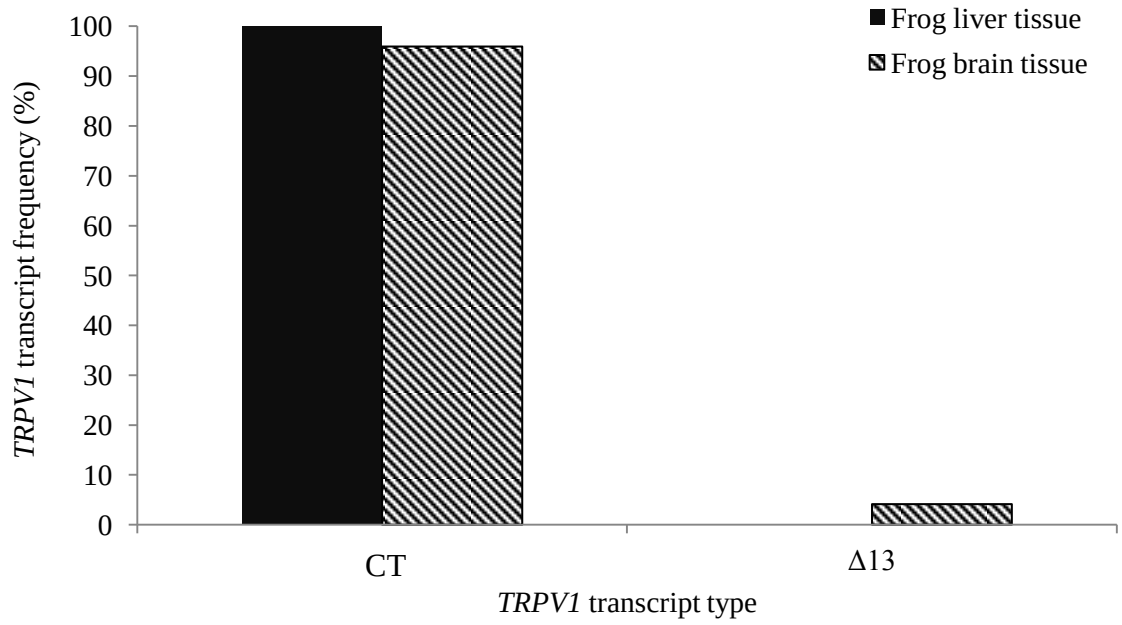

Figure 2.19 *TRPV1* CT and SV frequencies of the liver and brain tissues in the frog.

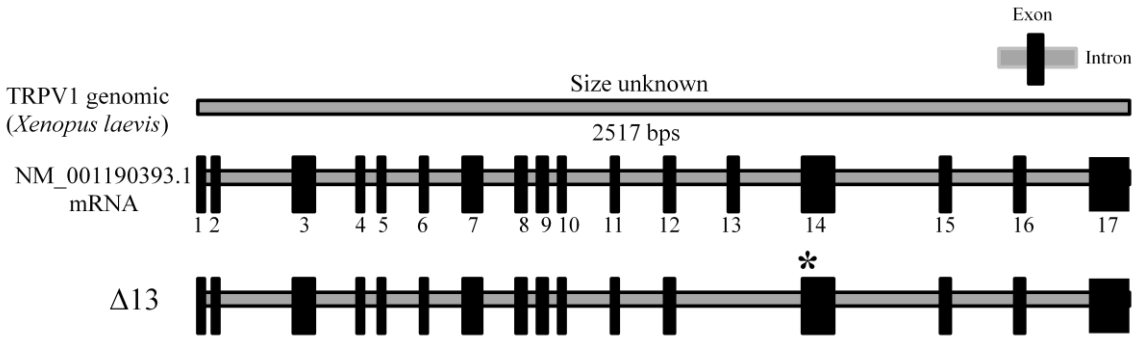

Figure 2.20 *TRPV1* transcript types characterized from the frog liver and brain tissues. * Indicates location of a PTC.

Table 2.7 *TRPV1* frequency of AS and transcript diversity in frog the liver and brain.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity (#) in tissue type
Liver	Canonical transcript	112	0	1
	Splice variant	0		
Brain	Canonical transcript	91	4.1	2
	Splice variant	4		

***TRPV1* and *POLB* SV frequencies in the frog liver and brain tissues**

There is a statistically significant difference between *POLB* SV frequencies and *TRPV1* SV frequencies in the liver of the frog (x^2: 1.8702, d.f=1, p=0.1714, α=0.05) (Figure 2.21). In the frog brain there is a statistically significant difference between *POLB* SV frequencies and *TRPV1* SV frequencies (x^2: 1.3393, d.f=1, p=0.2472, α=0.05) (Figure 2.21).

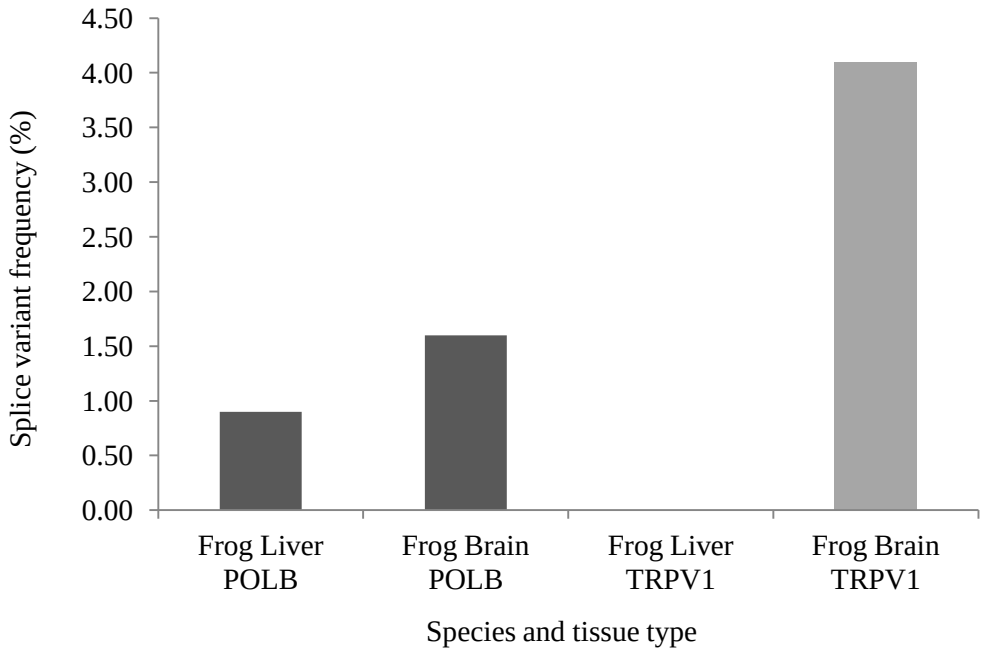

Figure 2.21 SV frequencies of *POLB* and *TRPV1* in frog liver and brain.

AS of *POLB* in bearded dragon tissues

POLB cloned transcripts were successfully amplified from bearded dragon liver and brain (Figure 2.22, lane 4 Figure 2.23 lane 4). The resulting size of the PCR product was approximately 1006 bps (Table 1.1). A total of 99 and 114 *POLB* transcripts were characterized from the frog liver and brain, respectively. A partial 1,006 bps of the canonical bearded dragon *POLB* transcript was cloned and characterized from this data (NCBI#: JN620208.1).

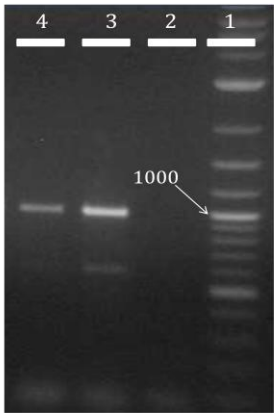

Figure 2.22 *POLB* PCR products amplified from bearded dragon liver tissue. Lane 1 shows the DNA size marker. Lane 2 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lanes 3 and 4 show duplicate successful amplification of *POLB* from bearded dragon liver cDNA.

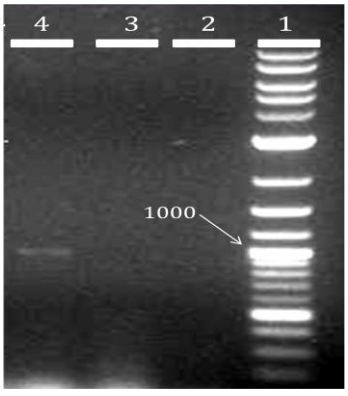

Figure 2.23 *POLB* PCR products amplified from bearded dragon brain tissue. Lane 1 shows the DNA size marker. Lanes 2 and 3 show lack of amplification from bearded dragon liver cDNA. Lane 4 shows successful amplification of POLB from bearded dragon brain cDNA.

AS of *POLB* in bearded dragon liver tissue

Of the 99 *POLB* transcripts characterized from the bearded dragon liver, two were SVs. The *POLB* SV frequency in the bearded dragon liver was 2.0% with a transcript diversity of three (Table 2.8). The first type of SV was Δ2 which occurred with a frequency of 1.0% (1/99) (Figure 2.24 and 2.25). Skip of *POLB* exon 2 resulted in an in frame PTC in exon three. The other SV characterized as a complete intron seven retention (ΣI7) (Figure 2.25). This SV occurred at frequency of 1.0% (1/99) and contained a PTC in the retained intronic sequence (Figure 2.24).

AS of *POLB* in bearded dragon brain tissue

Of the 114 *POLB* transcripts characterized the bearded dragon brain 11 SVs were observed. The *POLB* SV frequency in the bearded dragon brain was 9.6% with a transcript diversity of two (Table 2.8). This represented the highest *POLB* SV frequency in tissues from all non-mammalian species investigated. The one type of SV observed was Δ2 and it occurred at a frequency of 9.6% (11/114) (Figure 2.24 and 2.25). There was a statistically significant difference found in *POLB* SV frequency between bearded dragon liver and brain tissues (x^2=0.5, d.f=1, p=0.019, α=0.05).

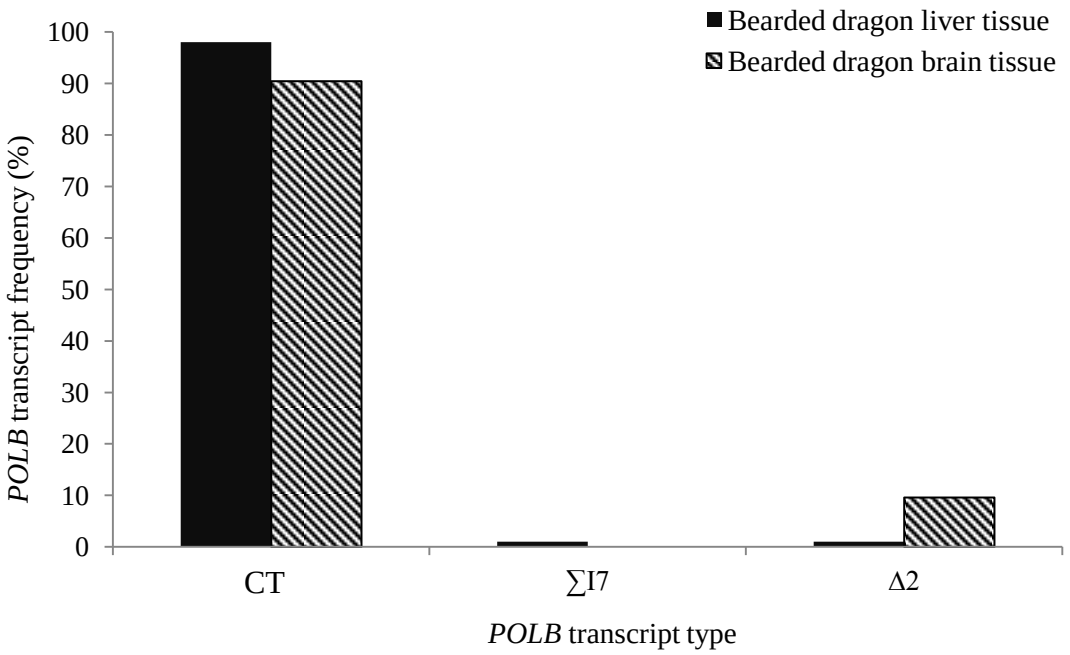

Figure 2.24 *POLB* CT and SV frequencies of the bearded dragon liver and brain tissues.

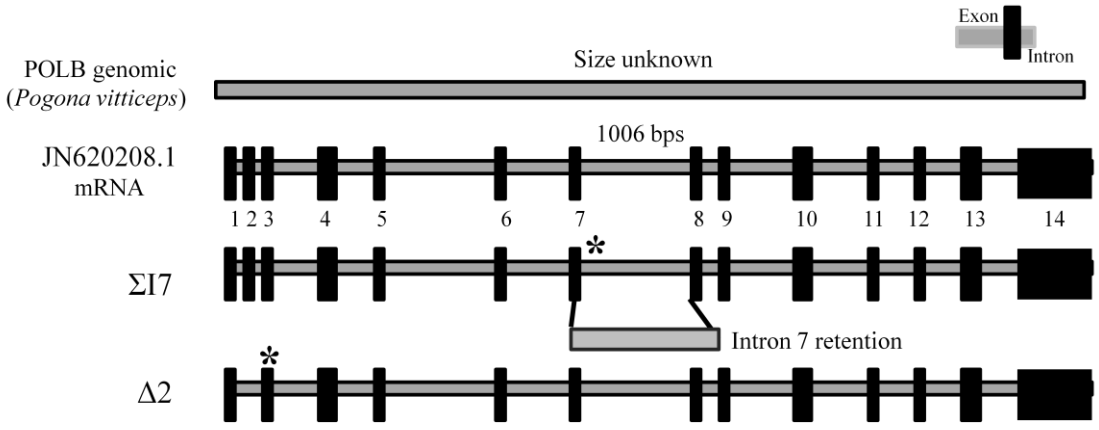

Figure 2.25 *POLB* transcript types characterized from the liver and brain tissues of the bearded dragon. * Indicates location of a PTC.

Table 2.8 *POLB* frequency of AS and transcript diversity in bearded dragon liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	97	2.0	3
	Splice variant	2		
Brain	Canonical transcript	103	9.6	2
	Splice variant	11		

AS of *POLB* and lifespan in non-mammalian tissues

The relationship between *POLB* SV frequency, transcript diversity and lifespan is shown in Figures 2.26 and 2.27. In the liver tissues of non-mammalian species AS of *POLB* did not correlate well with log lifespan (SV frequency r^2=0.0512, n.s; diversity r^2=0.0255, n.s) (Figure 2.26). In the brain tissues of non-mammalian species AS of *POLB* AS did not correlate well with log lifespan (SV frequency r^2=0.0255, n.s; diversity r^2=0.0255, n.s) (Figure 2.27).

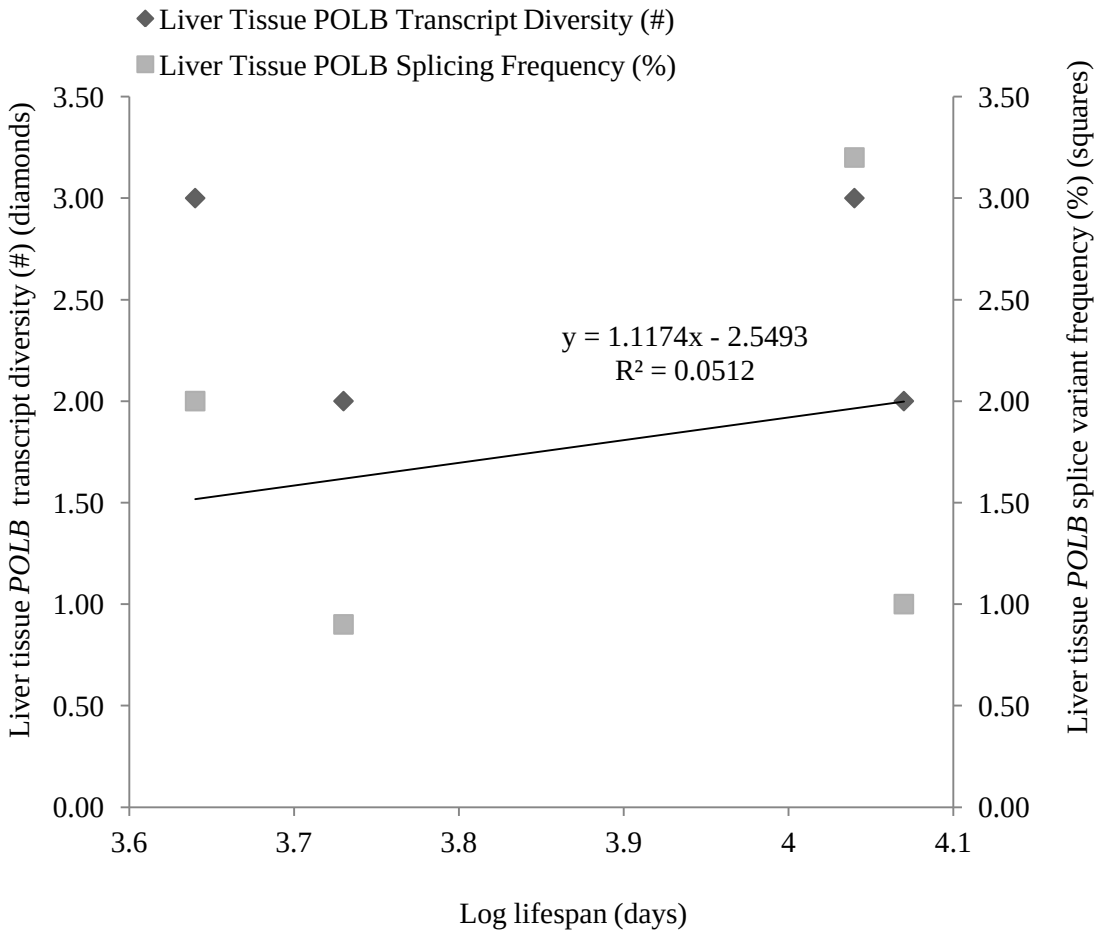

Figure 2.26 *POLB* SV frequency and transcript diversity with lifespan in liver tissues of non-mammalian species.

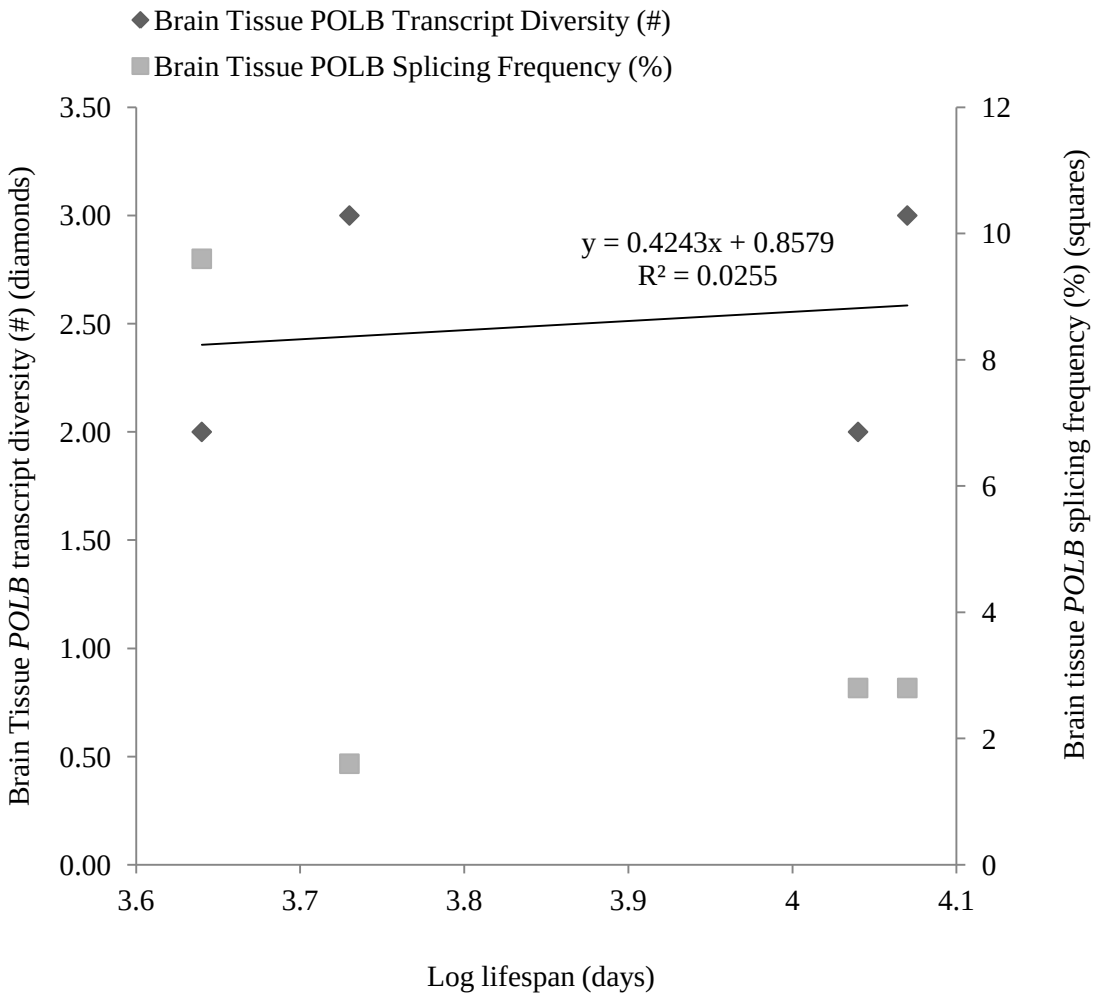

Figure 2.27 *POLB* SV frequency and transcript diversity with lifespan in brain tissues of non-mammalian species.

DISCUSSION

Although AS has been studied extensively in mammals, far fewer studies have focused on the phenomenon in non-mammals (Eichner et al., 2011; Schindler et al., 2007)**.** The majority of non-mammalian studies have focused on the AS of a single gene within one species. In cases of large scale EST or microarray studies the AS rates of certain genes have been reported for a few non-mammals, such as the domestic chicken (Kim et al., 2007, Artamonova et al., 2007). *POLB* is a highly conserved gene among vertebrate species (Uchiyama et al., 2009). In mammals there is an 87% conservation of *POLB* at the nucleotide level between rats and humans (Skandalis et al., 2010). In non-mammals, *POLB* mRNA sequence is also conserved. For example, there is 75% nucleotide level conservation between the domestic chicken and the African clawed frog.

Skandalis et al. (2010) revealed that *POLB* SV frequency varies, ranging from 12% to 57%, in the primary untransformed fibroblast cell lines of several primate species. The *POLB* splicing patterns in non-mammals were unknown prior to the work in this thesis. This chapter presents the analysis of data collected from individuals of three non-mammalian classes consisting of four orders: the chicken (Aves, Galliformes), glaucous-winged gull (Aves, Charadriiforms), African clawed frog (Amphibia, Anura), and bearded dragon (Reptilia, Squamata). Investigation in these species revealed that in comparison to mammals, *POLB* SV frequency and transcript diversity are low ($\leq$3.2%) in the liver and brain tissues in the majority of these non-mammalian species.

POLB SV frequency

The lowest *POLB* SV frequency was observed in the frog liver at 0.9%. The frog brain showed similar *POLB* SV frequency, at 1.6%. There is no indication of tissue specific AS of *POLB* in these two frog tissues. This is the first demonstration of the AS of *POLB* in a frog species and therefore, whether these frequencies are similar to those present in other amphibians is unknown at this time. Venables et al. (2012) reported the AS of other genes, for example, RNA-binding splicing factor (RBFOX), are found to alternatively splice in a tissue specific manner in the muscle and heart of the frog. However, the SV frequency of these AS events was not reported.

Similar *POLB* SV frequencies seen in frog tissues were observed in gull liver and brain tissues. These tissues did show a statistically significant difference in *POLB* SV frequencies; however this could be due to the number of transcripts analysed for each tissue. This may not represent a significant biological difference. To date there has been no investigation of the AS of any gene in the glaucous-winged gull other than the work presented in this chapter; however the *POLB* AS patterns of the gull were compared to those of another avian species, the chicken.

Two avian species were investigated in this work because of the possible influence of wide spread domestication of the chicken. The *POLB* SV frequencies were found to be higher in the chicken liver and brain tissues, 3.2% and 2.8%, respectively, compared to the gull tissues. There was no tissue specific splicing of *POLB* observed between the chicken liver tissue compared to the chicken brain tissue. A chi-square test between the *POLB* SV frequencies in the brain tissues of the two avian species showed no statistically significant difference. Interestingly, the liver tissues did show a

statistically significant difference for *POLB* SV frequency between the avian species. This result could suggest that widespread domestication of the chicken has affected the AS of certain genes in a tissue specific manner. While a recent study showed that other genes in the chicken alternatively splice in a highly tissue specific manner however the study did not comment on whether this could have been a result of domestication (Terenzi et al., 2010). Other studies have shown that the chicken has a whole organism AS rate of 42% with no suggestion of tissue specific AS events (Kim et al., 2007; Artamonova et al., 2007). These studies do not report the frequency at which the alternatively spliced genes occur. The results presented in this chapter show that *POLB* SV frequency in the chicken is similar to other non-mammalian vertebrates, and that these frequencies are low compared to primate species.

The highest *POLB* SV frequency of the four non-mammals was found in the bearded dragon brain, which occurred at 9.6%. This was 7.4% higher than the SV frequency in the bearded dragon liver. There was a statistically significant difference between the POLB SV frequency between the liver and the brain tissue of the bearded dragon. Chen et al. (1997) showed that other genes in reptiles do demonstrate tissue specific AS.

With the exception of the bearded dragon, the non-mammalian tissues were obtained from individuals in the young adult stage of life. Tollervey et al. (2011) reported a general increase in AS with age in humans. The bearded dragon liver and brain tissues were obtained from an individual of 11 years of age while the recorded lifespan of a bearded dragon is 12 years (Reusch et al., 2009). The higher *POLB* SV frequency in this species as compared to the other non-mammals could be a consequence of advanced age.

Taken together, the *POLB* SV frequencies in these species are lower than reported in primates by Skandalis et al. (2010). There appears no tissue specific AS of *POLB* between liver tissues compared with brain tissues in the majority of non-mammalian vertebrates.

POLB transcript diversity

The *POLB* transcript diversity was lower in all non-mammals as compared to primates (Skandalis et al., 2010). No more than one to two types of *POLB* SVs were observed in these species compared to the ≥6 - 14 types of SVs in reported for primate species (Skandalis et al., 2010). This result suggests that the chicken, gull, frog and bearded dragon produce fewer different types of *POLB* SVs than primates.

A single type of *POLB* AS event appears to be conserved: the skip of exon two. The Δ2 *POLB* SV was observed in some form in all non-mammals with the exception of the chicken. In two cases, the gull and the frog, the skip of exon two was always involved in a multiple exon skip which included exon three. Additionally, the skipping of exon two AS events in *POLB* SVs was only observed in brain tissue. In humans, Δ2 has been observed in bladder tissues, gastric tissues and blood samples from healthy individuals as well as various cancer cell types (Simonelli et al., 2009; Thompson et al., 2002). Skandalis et al. (2010) suggest that the Δ2 SV is evolutionarily conserved among primate species. The data in this chapter presents the first evidence of the Δ2 *POLB* variant in non-mammalian vertebrates, which suggests that this SV may be conserved in the majority of vertebrates. This SV does not produce a functional protein and has been proposed to act as a dominant negative regulator of normal POLB activity (Thompson et al., 2002). Recently, it has been strongly suggested that the Δ2 variant functions as a non-

coding RNA that plays a role as a post-transcriptional regulator in humans (Simonelli et al., 2009). Whether exon two skip in *POLB* SVs has similar functions in the brain of the species investigated in this chapter as those suggested in humans is currently unknown.

Intron retention AS events

Intron retention AS events in *POLB* SVs, either full or partial, were observed in all species investigated in this chapter. In general, vertebrates have short exons compared to introns (reviewed by Keren et al., 2010). The number of large introns, >50 kbs, has been shown to increase from non-mammals to mammals (Shepard et al., 2009).

The PCR conditions and Taq polymerase used in this thesis work are generally capable of amplifying products <2.0kb (Evans 2008; Barnes 1994). Introns that are short in length may not have been spliced out of the mRNA at the time of the total RNA isolation and therefore may have been observed in the cloned transcripts. For example, intron one of *POLB* in the chicken is 71 bps. This small intron may have been detected simply because it had not yet been spliced out of the mRNA during co-transcriptional splicing, and not necessarily because it was an intron retention AS event. To investigate whether the intron retention event observed in the chicken *POLB* AS patterns was due to the small intron length, POLB transcripts could be characterized from isolated nuclear and cytoplasmic RNA fractions and compared. If the intron was present in a *POLB* transcript detected from cytoplasmic RNA, this would suggest an intron retention AS event.

The intron retention AS events observed in non-mammalian *POLB* SV may be a result of the secondary structures of the polypyrimidine binding tract (PPT) of the intron. Recently, Vargas et al. (2011) found evidence that indicated the spliceosome assembled

at each intron independently, in turn catalyzing the splicing reaction depending on the unique kinetics of each intron. Furthermore, authors observed that when an intron's PPT formed strong secondary structures splicing of the intron was delayed until after transcription was complete even though all other introns were co-transcriptionally spliced out (Vargas et al., 2011). This delay resulted in an accumulation of incompletely spliced mRNAs in the nucleus. The intron retention AS events seen in non-mammalian *POLB* transcripts could be a result of the secondary structures formed by the intron's PTT and this unique circumstance of uncoupled splicing. Unfortunately, since the *POLB* intronic sequences are unknown for three out of the four species in this chapter an analysis of possible intron secondary structures across these species was not possible.

POLB **AS patterns versus** ***TRPV1*** **patterns**

Chicken and frog *POLB* AS patterns were compared to the patterns of a second gene, *TRPV1*, to investigate whether the observed *POLB* splicing patterns occurred due to a systemic (overall) increase in AS in a particular tissue type. Both the chicken liver and the brain showed statistically significant differences in the SV frequencies of *POLB* and *TRPV1*. *TRPV1* SV frequency in the chicken liver was increased compared to the *POLB* SV frequency, at 78.0% and 3.2%, respectively. This result suggests that the chicken liver is not subject to an overall increase in AS. If the liver of the chicken had increased AS frequencies for all genes as compared to the brain then high SV frequencies of both *POLB* and *TRPV1* would be observed. This was not reflected in the data presented in this chapter. The brain and liver tissue of the chicken showed similar results; both tissues had higher SV frequency of *TRPV1* compared to *POLB*.

The frog liver and brain tissues also showed statistically significant differences in *POLB* and *TRPV1* SV frequencies. In the frog brain tissue, *POLB* and *TRPV1* SV frequencies were higher than in the liver. No *TRPV1* SVs were detected in the frog liver whereas *POLB* SVs were observed. These data may suggest that AS of all genes is higher in the brain compared to the liver for this species. To further investigate this possibility, future studies should examine the AS patterns of various genes in these two tissues in non-mammals. The genes investigated in those studies should be a combination of DNA repair genes, non-DNA repair genes and additional genes of cation channels. If the AS patterns of all genes are increased in brain tissue as compared to liver tissue this would provide further data to suggest the presence of an overall increase in AS in particular tissues of this species.

Chickens are insensitive to capsaicin because the avian *TRPV1* channel fails to be activated by capsaicin, although it is still sensitive to heat and protons (Jordt et al., 2002). Jordt et al. (2002) reported that the chicken TRPV1 channel differed in ~8 amino acids located in the third transmembrane domain compared to humans. The difference in this domain was found to be responsible for the insensitivity to capsaicin in chickens. Insensitivity to capsaicin has also been reported in reptiles and amphibians (Szallasi et al., 1999; Hawkins et al., 1991). Functional TRPV1 proteins form as tetramers and exist as homo- or heteromultimers (Jordt et al., 2002). This tetramer protein structure may indicate that the inclusion of one of more *TRPV1* SVs could affect the function of the channel. Furthermore, *TRPV1* SVs in humans have been reported to form TRPV1 channels which are not activated by capsaicin or protons (Lu et al., 2005). Perhaps with less pressure to maintain conservation of functional TRPV1 channels for detection of

capsaicin avian *TRPV1* mRNA is more likely to undergo AS without possible detrimental results from unproductive SVs. These data may also suggest that the AS of avian *TRPV1* is more likely to occur in certain tissue types based on the high frequency of *TRPV1* AS found in the liver as compared to the brain.

The correlation between the AS of *POLB* and lifespan in non-mammals

Finally, the AS of *POLB* in both tissue types for all four non-mammals investigated did not correlate well with lifespan. Unlike the strong positive correlation between *POLB* SV frequency reported in primates (Skandalis et al., 2010), non-mammalian vertebrates did not show a correlation between an increase in *POLB* SV frequency and lifespan. These data suggest that the mechanisms in primates which generate *POLB* SVs are different than ones used by non-mammalian vertebrates.

The results in this chapter agree with previous data found in primates that showed most *POLB* SV types are not conserved among species and contain PTCs (Skandalis et al., 2010). No single type of *POLB* SV was found throughout all non-mammalian species. However, the exon two skip event was present in SVs from three out of the four species in this chapter.

CONCLUSIONS

The data presented in this chapter do not support the hypothesis that a positive correlation between *POLB* SV frequency and lifespan exists for non-mammalian vertebrates. This was the first work done on the AS of *POLB* in non-mammals and it revealed that SV frequencies and transcript diversities of *POLB* were low in these species compared to primates. Interestingly, the exon two skip AS event appeared to be conserved in the majority of vertebrate species. Tissue specific *POLB* AS was observed in the gull and the bearded dragon.

Comparison with a second gene, *TRPV1* suggested that no increase or decrease of AS was present in specific tissues in the chicken whereas an increase in AS of both genes was observed in the frog brain compared to liver tissue. While non-mammalian vertebrates produced non-functional *POLB* SVs, similar to primates, the SVs occurred at lower frequencies and with less transcript diversity. Taken together, these results indicated that the mechanism for AS of *POLB* in primates that produced non functional *POLB* SVs at higher frequency in correlation with a longer lifespan in primates was not present in non-mammalian vertebrates.

CHAPTER THREE: Alternative splicing of *POLB* in mammalian tissues and cell lines

Goal of Chapter three

The major goal of chapter three was to compare the AS patterns of *POLB* in mammalian tissues to those observed in cell lines. Previous literature has shown that in untransformed human fibroblasts *POLB* SV frequency and diversity are 57.4% and 14, respectively (Skandalis et al., 2010). In conjunction with the AS patterns of *POLB* in humans and four other primate species, *P. pygmaeus*, *G. gorilla*, *P. troglodytes* and *M. fascicularis,* Skandalis et al. (2010) revealed a positive correlation between *POLB* SV frequency and lifespan. What remains to be determined is if the AS patterns of *POLB* are consistent in mammalian tissues versus untransformed fibroblast cell lines. To address this, the *POLB* SV frequency and transcript diversity was compared in human liver and brain tissues to human cell lines. Furthermore, a comparison was done between mammalian and non-mammalian liver and brain tissues to examine whether a positive correlation between *POLB* SV frequency and lifespan exists in vertebrate species.

Similar to chapter two this chapter examined tissue specific AS patterns of *POLB* as well as using a second gene, *TRPV1*, to investigate possible tissue specific or gene specific changes in AS in specific tissue types.

This chapter will explore three major hypotheses. First, I hypothesized that the AS patterns of *POLB* are consistent in human fibroblast cell lines to patterns observed in RNA isolated from human liver and brain tissues. Secondly, I hypothesized that the *POLB* SV frequency in mammalian tissues and primate cell lines, positively correlate with species lifespan. Finally, I hypothesized that a positive correlation between *POLB* SV frequency and lifespan will be seen in all vertebrate species studied here.

INTRODUCTION

Numerous approaches to explore AS have been employed which utilize *in vivo*, *in vitro* and computational bioinformatics systems (Florea 2005). There are advantages and disadvantages to each approach which can influence the AS patterns observed in each system. For example, EST studies which evaluate AS can only use the data available in databases, such as GeneBank and the Alternative Splicing Database. This could result either in an over or underestimation of AS frequencies, SV diversities, the types and AS events and their frequencies (Martin et al., 2011; Filichkin et al., 2010; Koralewski et al., 2010; Gaidatzis et al., 2009). This matter has initiated comparative studies which evaluate AS data collected from different systems (Wang et al., 2008; Möröy T. et al., 2007; Kim et al., 2006; Helfman et al., 1998). This chapter focuses on comparing the AS patterns of *POLB* in mammalian tissues and cell lines.

Generally, AS has been studied best using cell lines (Li et al., 2006). One reason for this trend is that cell lines generally represent a single cell type which provides a higher degree of reproducibility (Fedoroff et al., 2001). This is difficult to obtain through use of tissues, which often contain many cell types. Experimental conditions during cell culture can be tightly controlled allowing little to no variation in energy availability, growth factor levels and oxygen exposure. Another advantage to cell lines is a considerable decrease in cost. The cost of using animals for experimental work is very high in comparison to cell lines. In addition to the high cost of maintaining animals for experiments other concerns, such as ethical issues, are not applicable when working with cell lines. Tissue collection can be costly even if whole animals are not required. For instance, obtaining tissue samples from endangered species is sometimes impossible, but

acquiring cell lines is not out of the question. The living conditions and experiences of the animal prior to the point of sampling may also affect the AS patterns of the organism.

There are proposed limitations to using cell lines for AS studies. For example, cells rarely exhibit *in vivo* phenotypes of similar cell types (Li et al., 2006; Kemp et al. 2005). Lu et al. (2006) reported unique AS events were seen in the same cell line when grown under different conditions (Li et al., 2006). In addition, Li et al. (2006) found that AS patterns in tumour tissues more closely resembled patterns seen in Matrigel-cultured cells as compared to flat dish-cultured cells. This was suggested to occur because Matrigel-cultured cells contain a media which closely resembles the complex extracellular environment found in *in vivo,* whereas flat-dish cultured cells do not (Li et al., 2006; Moghe et al., 1997).

Kemp et al. (2005) described differences between AS events in cell culture versus tissues from transgenic mice. In this study, cell culture provided an incomplete picture of transcriptional regulation as compared to eight tissue types (Kemp et al., 2005). Thompson et al. (2002) showed the *POLB* AS had the similar SV frequencies (~ 55%) and transcript diversities (~ 6) in bladder cancer tissues versus bladder cancer cell lines. Thompson et al. (2002) was the only study to date which investigated the AS patterns of *POLB* in the cell lines and tissues, however testing was only conducted in bladder cancer samples with ≤31 cloned transcripts characterized.

POLB AS data collected from cell lines have elucidated important features about the AS of this gene such as SV types, SV frequencies, and suggested function for several *POLB* SVs (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002). Skandalis et al., (2010, 2004) and Disher et al. (2007)

showed *POLB* AS patterns in human untransformed fibroblast cell lines demonstrated numerous *POLB* SV types and high *POLB* SV frequencies (44.2%-57.4%) compared to rat fibroblasts (~1%-3%). It is important to investigate whether similar *POLB* AS patterns are present in human liver and brain tissues. Similar AS patterns will help support the hypothesis that SVs seen in cell lines are representative of the organism as a whole. This issue will be addressed by comparing the AS patterns *POLB* reported in cell lines by Skandalis et al. (2010) to liver and brain RNA isolated from human tissues.

A second mammal, the rabbit, was used to determine whether *POLB* SV frequency and transcript diversity correlate to lifespan in other mammalian species, as previously reported for primate species (Skandalis et al., 2010). Skandalis et al. (2010) has previously suggested that this correlation does exist for mammals in general. The *POLB* AS data from rabbit tissues will contribute to a growing mammalian data collection.

Finally, mammalian *POLB* AS patterns from this chapter will be compared to the non-mammalian data reported in chapter two to test the hypothesis that a positive correlation between the AS of *POLB* and lifespan exists in vertebrates.

RESULTS

AS of *POLB* in rabbit tissues

POLB cloned transcripts were successfully amplified from the rabbit liver and brain (Figure 3.1, lane 2 and Figure 3.2 lane 10). The resulting size of the PCR product for the rabbit tissues were approximately 965 bps (Table 1.1). A total of 66 and 50 *POLB* transcripts were characterized from the rabbit liver and brain, respectively.

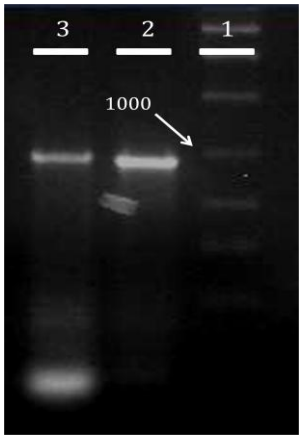

Figure 3.1 *POLB* PCR products amplified from rabbit liver. Lane 1 shows the DNA size marker. Lane 2 shows successful amplification of *POLB* from cDNA from the rabbit liver using POLBuniF x R. Lane 3 shows successful amplification of *POLB* from cDNA from the rabbit liver using DEGPOLBF x R.

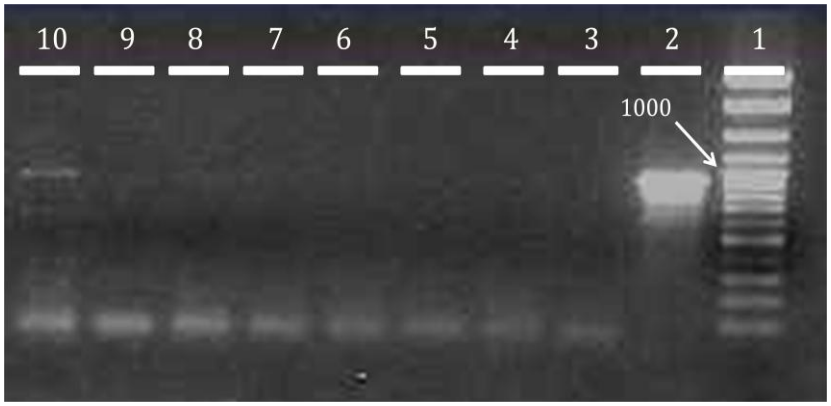

Figure 3.2 *POLB* PCR products amplified from rabbit brain tissue. Lane 1 shows the DNA size marker. Lane 2 contains a positive control, in which rabbit liver colony template was used. Lane 10 shows successful amplification of *POLB* from cDNA from the rabbit brain. Lanes 3-9 show unsuccessful amplification of *POLB* from other cDNA from the rabbit liver.

AS of *POLB* in rabbit liver tissue

Of the 66 *POLB* transcripts characterized from the rabbit liver 13 were SVs. *POLB* SV frequency in the rabbit liver was found to be 19.7% with 8 unique transcripts (Table 3.3). The non-canonical transcript with the highest frequency was retention of 78 bps of undetermined sequence which occurred at a frequency of 7.6% (4/66) (Figure 3.3). This retention was determined through restriction digest with EcoRI which showed approximately 78 extra bps (according to restriction digest gel electrophoresis) in the second half of the gene (exons 8-14). Sequencing of this transcript was unsuccessful despite repeated attempts.

Two types of SVs, Δ4-6 and Δ9, were found with a frequency of 3% (2/66) each (Figure 3.2 and 3.3). The Δ4-6 transcript contains a premature termination codon (PTC) in exon eight created by a multiple exon skip. The Δ9 *POLB* SV contained a PTC in exon ten. The remaining *POLB* SVs observed all occurred at a frequency of 1.52% (1/66) (Figure 3.3). These SVs were Δ4,5,9, Δ4-10, Δ11 and ∑pI9$_{9}$,Δ12,13 (Figure 3.4). The Δ4,5,9 SV is 207 bps less than the CT. Multiple exon skipping of exons four and five caused a PTC in exon six. The Δ4-10 *POLB* SV is 435 bps less than the CT. The multiple exons four through ten skip did not disrupt the reading frame and there are no PTCs included in the mRNA due to the seven skipped exons. The Δ11 SV is missing 87 bps due to the exon 11 skip, compared to the CT. The exon 11 skip did not cause a PTC in the transcript. Finally, a complex partial intron retention and multiple exon skip event was observed in the ∑pI9$_{9}$Δ12,13 *POLB* SV (Figure 3.4). The retention of nine bps of intron nine in addition with the exon 12 and 13 skipping did not cause a PTC in this transcript.

AS of *POLB* in rabbit brain tissue

Of the 50 *POLB* transcripts characterized from the rabbit brain ten were SVs. The *POLB* SV frequency in the rabbit liver was 20% with a transcript diversity of eight (Table 3.1). The SVs which occurred with the highest frequency, at 4.0% (2/50), were Δ2,9, Δ2,12, and Δ4,5 (Figures 3.3 and 3.4). The Δ2,9 and Δ2,12 SVs both contained PTCs in exon 3 due to exon 2 skip. The Δ4,5 *POLB* SV is 134 bps compared to the CT sequence due to the multiple exon four and five skip. This splicing event causes a PTC in exon six.

The remaining *POLB* SVs characterized in the rabbit brain occurred at frequencies of 2.0% (1/50) (Figure 3.3). These SVs were Δ2,11, Δ9, Δ9,11 and Δ12,13 (Figure 3.4). The skipping of exon nine caused a PTC in affected transcripts in exon ten. There was no statistically significant difference in *POLB* SV frequencies between the rabbit liver and brain (x^2=0.0016, d.f=1, p=0.9677, α=0.05).

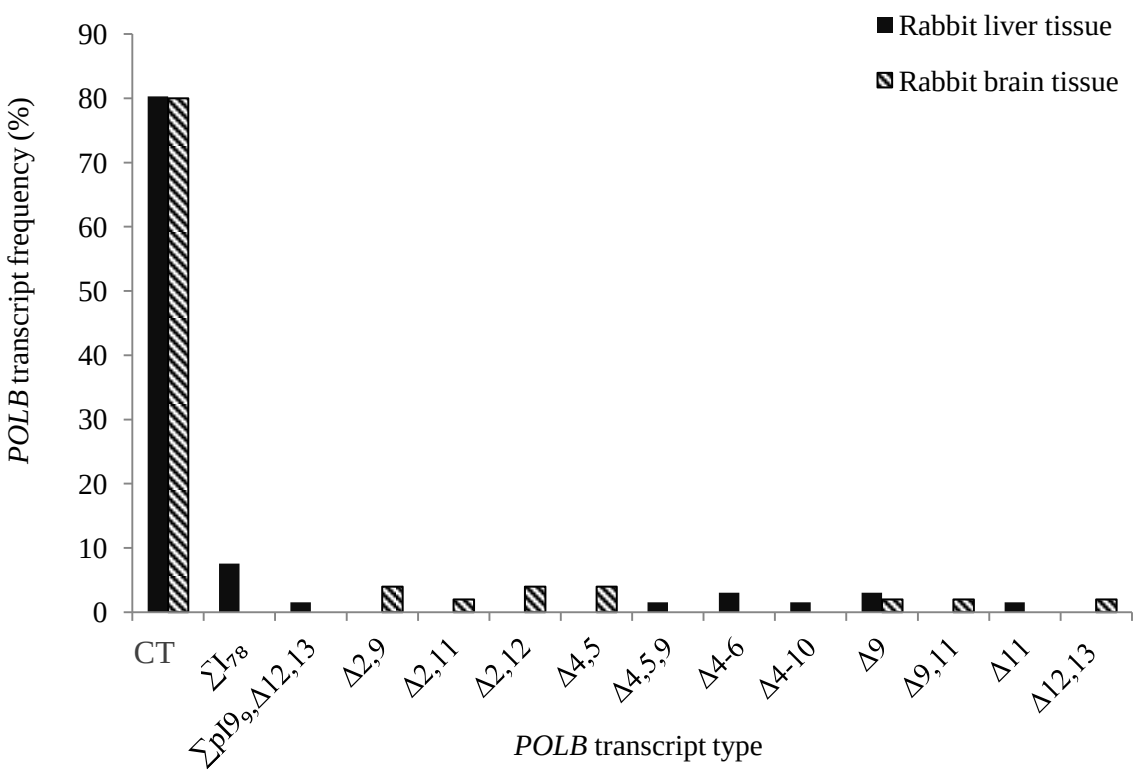

Figure 3.3 *POLB* CT and SV frequencies of the liver and brain tissues in the rabbit.

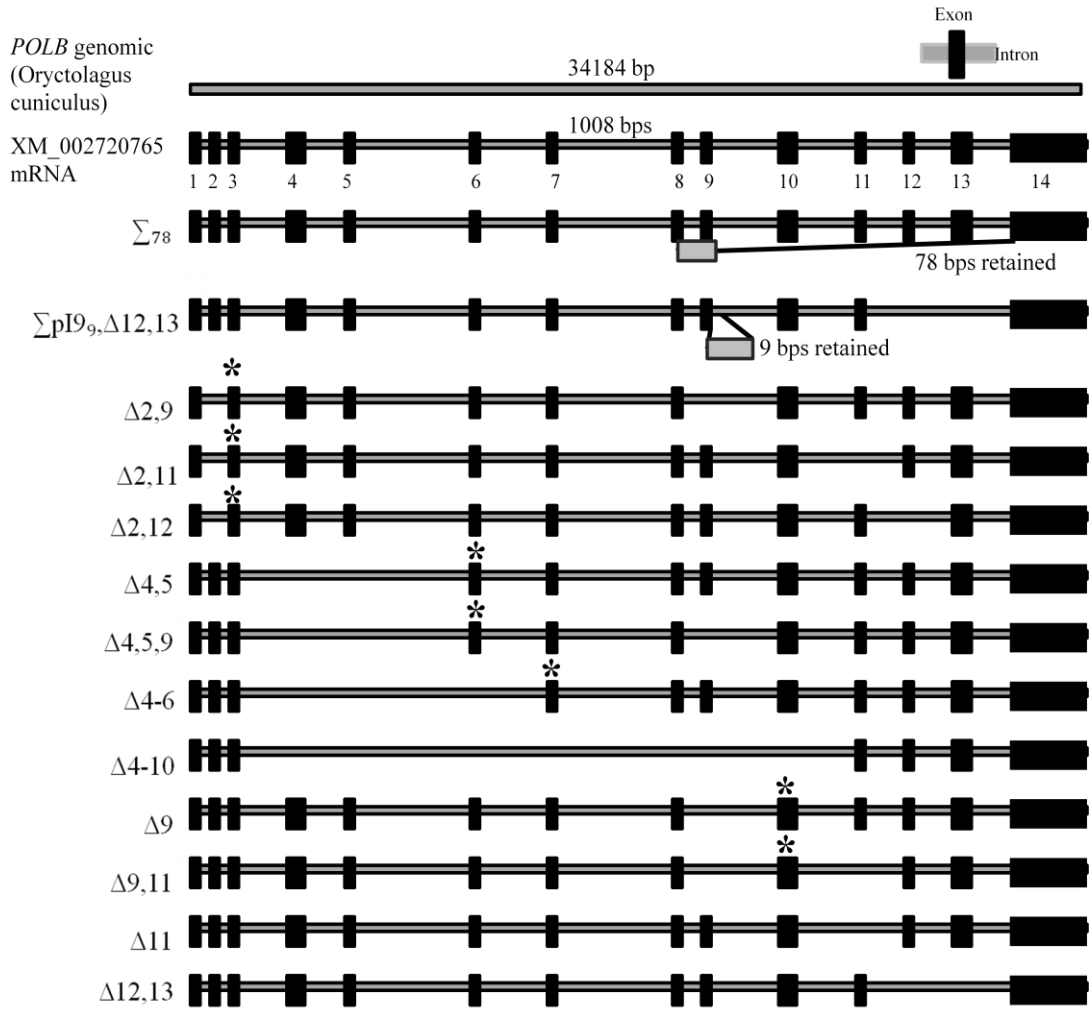

Figure 3.4 *POLB* transcript types characterized from the liver and brain tissues of the rabbit. * Indicates location of a PTC.

Table 3.1 *POLB* frequency of AS and transcript diversity in rabbit liver and brain tissues.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	53	19.7	8
	Splice variant	13		
Brain	Canonical transcript	40	20.0	8
	Splice variant	10		

AS of *POLB* in human tissues

POLB cloned transcripts were successfully amplified from human liver and brain total RNA (Figure 3.5, lane 8 and Figure 3.6 lanes 2-5). The resulting size of the PCR product for the human liver was approximately 1121 bps (Table 1.1). A total of 153 and 96 *POLB* transcripts were characterized from the human liver and brain, respectively.

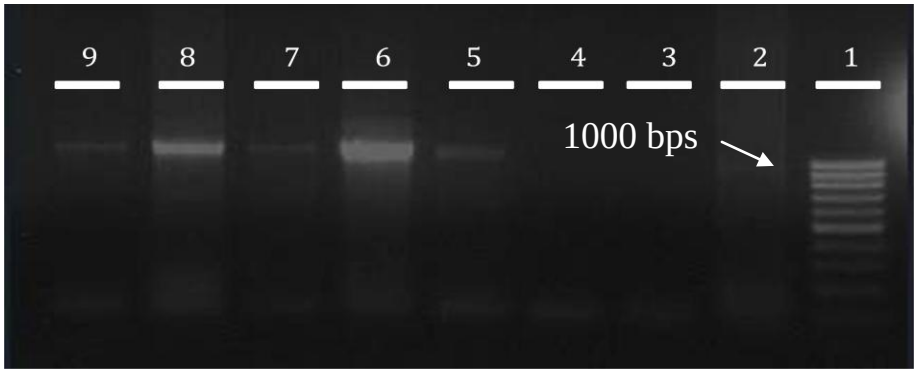

Figure 3.5 *POLB* PCR products amplified from human liver. Lane 1 shows the DNA size marker. Lane 2 shows unsuccessful *POLB* amplification from human liver cDNA using POLBuniF x R. Lane 3 contains a negative control, in which no cDNA was added to the reaction mixture, but water was. Lane 4 contains a second negative3 control, in which rabbit liver cDNA was used with human *POLB* primers, POLBfor1 x POLBrev1. Lanes 5 and 8 show successful amplification of *POLB* from human liver cDNA. Lane 6 shows successful amplification of *POLB* from human breast tumor cDNA, which was not used in this work. Lane 7 and 9 show successful amplification of *POLB* from human fibroblast cell lines, as a positive control.

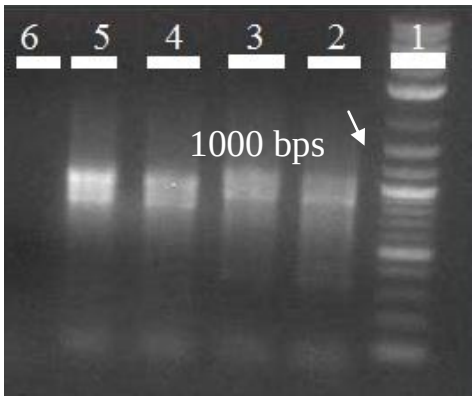

Figure 3.6 *POLB* PCR products amplified from human brain. Lane 1 shows the DNA size marker. Lanes 2-5 show *POLB* PCR product from the human brain tissue. Lane 6 contains a negative control, where water was used in place of template.

AS of *POLB* in human liver tissue

Of the 153 *POLB* transcripts characterized from the human liver 100 were SVs. The *POLB* SV frequency in the human liver was found to be 65.4% (Table 3.2). The *POLB* transcript diversity of the human liver was 18 (Table 3.2). The 17 types of SVs that were observed in this tissue showed simple and complex splicing events, including partial intron retention, single exon skipping, multiple exon skipping and partial intron retention along with single/multiple exon skipping.

The *POLB* SV which was observed with the highest frequency in the human liver was the Δ2 SV (Figure 3.7). This SV contained a PTC in exon 3 due to exon 2 skipping. This SV occurred with equal frequency with the CT at 34.4% (53/153) (Figure 3.9). The SV which occurred with the second highest frequency at 5.8% (9/153), was the *POLB* Δ11 SV (Figures 3.7 and 3.9). This SV did not contain a PTC as a result of exon skipping.

A single partial intron retention event was observed in the human liver. The partial intron retention event observed was the retention of 105 bps from intron six, referred to here as Σα (Figure 3.7). The inclusion of Σα does not disrupt the reading frame of the *POLB* transcript. This SV occurred at a frequency of 3.3% (5/153) (Figure 3.9).

Three other SVs were found all with a frequency of 3.3% (5/153): Δ2,4, Δ2,11, and Δ2,4-6 (Figures 3.7 and 3.9). Similar to all other *POLB* SVs with exon two skip and an intact exon three a PTC is created in exon three for all of these SVs.

Two *POLB* complex partial intron retention and single exon skip splicing events were observed in the human liver. The first was a SV which retained Σα and had exon two skip (Figure 3.7). This Σα, Δ2 SV had a PTC in exon three. The second complex

partial intron retention and single exon skip was the *POLB* SV Σα,Δ11 and it occurred at a frequency of 1.3% (2/153) (Figure 3.7 and 3.9). There was no PTC in this transcript as a result of the intron inclusion and exon skip.

Three *POLB* SVs were found to occur with a frequency of 1.3% (2/152); Δ2,5, Δ2,4-6,11, and Δ6,11 (Figures 3.7 and 3.9). The Δ2,5 and Δ2,4-6,11 SVs had PTCs in exon three due to exon two skip. Finally, the Δ6,11 SV contained a PTC in exon seven.

The final six *POLB* SVs observed in the human liver all occurred at a frequency of 0.65% (1/153) (Figure 3.9). These six SVs were Δ2,4,5, Δ2,4-6,9,11, Δ3-5,11, Δ4,5,11, Δ5-10 and Δ12,13 (Figure 3.7). A PTC is created in exon three in the *POLB* SVs Δ2,4,5 and Δ2,4-6,9,11. The Δ4,5,11 *POLB* SVs contained a PTC in exon six. Finally, the Δ3-5,11, Δ5-10 and Δ12,13 SVs did not contain any PTCs due to the multiple exon skipping splicing events.

AS of *POLB* in human brain tissue

Of the 96 *POLB* transcripts characterized from the human brain 88 were SVs. The *POLB* SV frequency in the human brain was found to be 91.7% (Table 3.2). The *POLB* transcript diversity of the human brain was 25 (Table 3.2). This tissue had the highest *POLB* SV frequency and diversity of any tissue investigated here in all species. Similar to the human liver, the types of splicing events seen in the human brain were composed of simple and complex intron retentions, single and multiple exon skipping.

Five of the *POLB* SV types found in the brain were previously observed in the human liver. These SVs included; Σα, Σα,Δ2, Δ2, Δ2,4-6,11, and Δ11 at frequencies of 1.0% (1/96), 2.1% (2/96), 4.2% (4/96), 4.2% (4/96), 3.1% (3/96), and 3.1% (3/96), respectively (Figures 3.8 and 3.9). A *POLB* intron retention splicing event was seen in the

human brain. This SV occurred at a frequency of 1.1% (1/96) and retained 18 bps from intron eight, 108 bps from intron ten and had multiple exon skipping of exons 2,4-7,12, and 13 (Figures 3.8 and 3.9). Due to exon two skip the $\sum p8_{18}$, $\sum p10_{108}$, Δ2,4-7,12,13 SV contained a PTC in exon three.

In addition to the SVs previously described in the human liver, ten SVs were found in the human brain which had exon two skipping with exon three intact. These transcripts and their frequencies were: Δ2,4-5,11-13 (1.0%) (1/96), Δ2,4-6 (16.7%) (16/96), Δ2,4-6,12,13 (1.0%) (1/96), Δ2,4-6,12 (3.1%) (3/96), Δ2,4-7,9 (2.1%) (2/96), Δ2,4-9,12 (7.3%) (7/96), Δ2,4-9,12,13 (6.3%) (6/96), Δ2,4-11 (3.1%) (3/96), Δ2,4-12 (2.1%) (2/96), and Δ2,4-13 (22.9%) (22/96) (Figures 3.8 and 3.9). A PTC was present in exon three for these SVs.

Six *POLB* SVs in the human brain had multiple exon skips in the middle of the transcripts starting with the skip of exon two. These SVs and their frequencies were: Δ2-4 (4.2%) (4/96), Δ2-5 (1.1%) (1/96), Δ2-7,9 (1.0%) (1/96), Δ2-9 (1.0%) (1/96), Δ2-9,11-13 (2.1%) (2/96) and Δ2-13 (1.0%) (1/96) (Figures 3.8 and 3.9). A PTC was created in the SVs Δ2-4, Δ2-5, and Δ2-7,9 in exons five, six, and 11, respectively. Despite the $\geq$8 exon skipping in the SVs Δ2-9, Δ2-9,11-13 and Δ2-13 no PTCs were present in these transcripts.

One *POLB* SV was found in the human brain which was also observed in the rabbit liver. This SV had exon 4-6 skipping and occurred at a frequency of 1.0% (1/96) (Figures 3.8 and 3.9). The Δ4-6 SV had a PTC in exon eight.

Finally, the last *POLB* SV found in the human brain was Δ11-13 and occurred at a frequency of 1.0% (1/96) (Figures 3.8 and 3.9). This SV did not contain any PTCs. There

was a statistically significant difference in *POLB* SV frequencies between human liver and brain (x^2=22.0711, d.f=1, p=>0.0001, α=0.05).

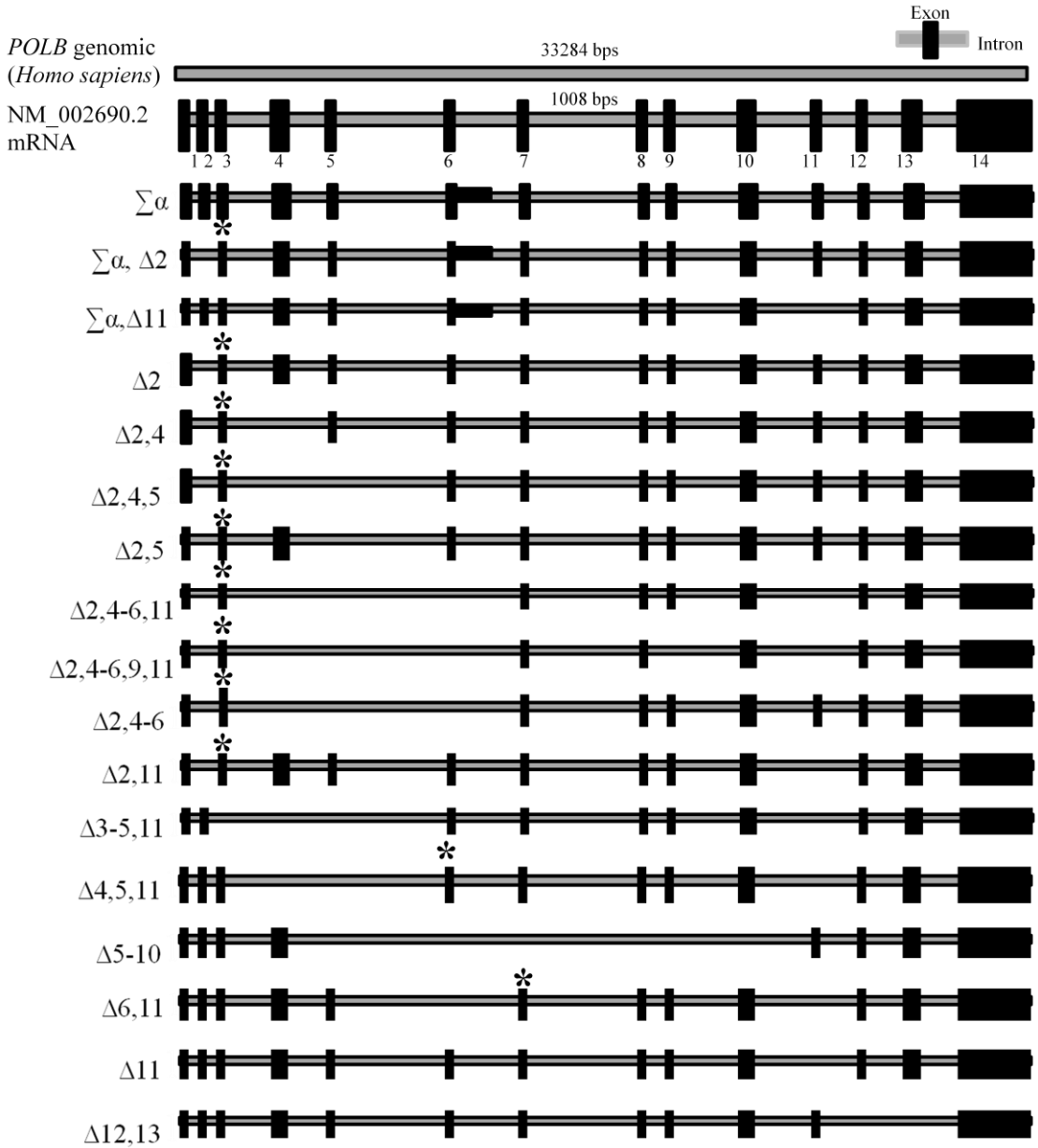

Figure 3.7 *POLB* transcript types characterized from the human liver. * Indicates location of a PTC.

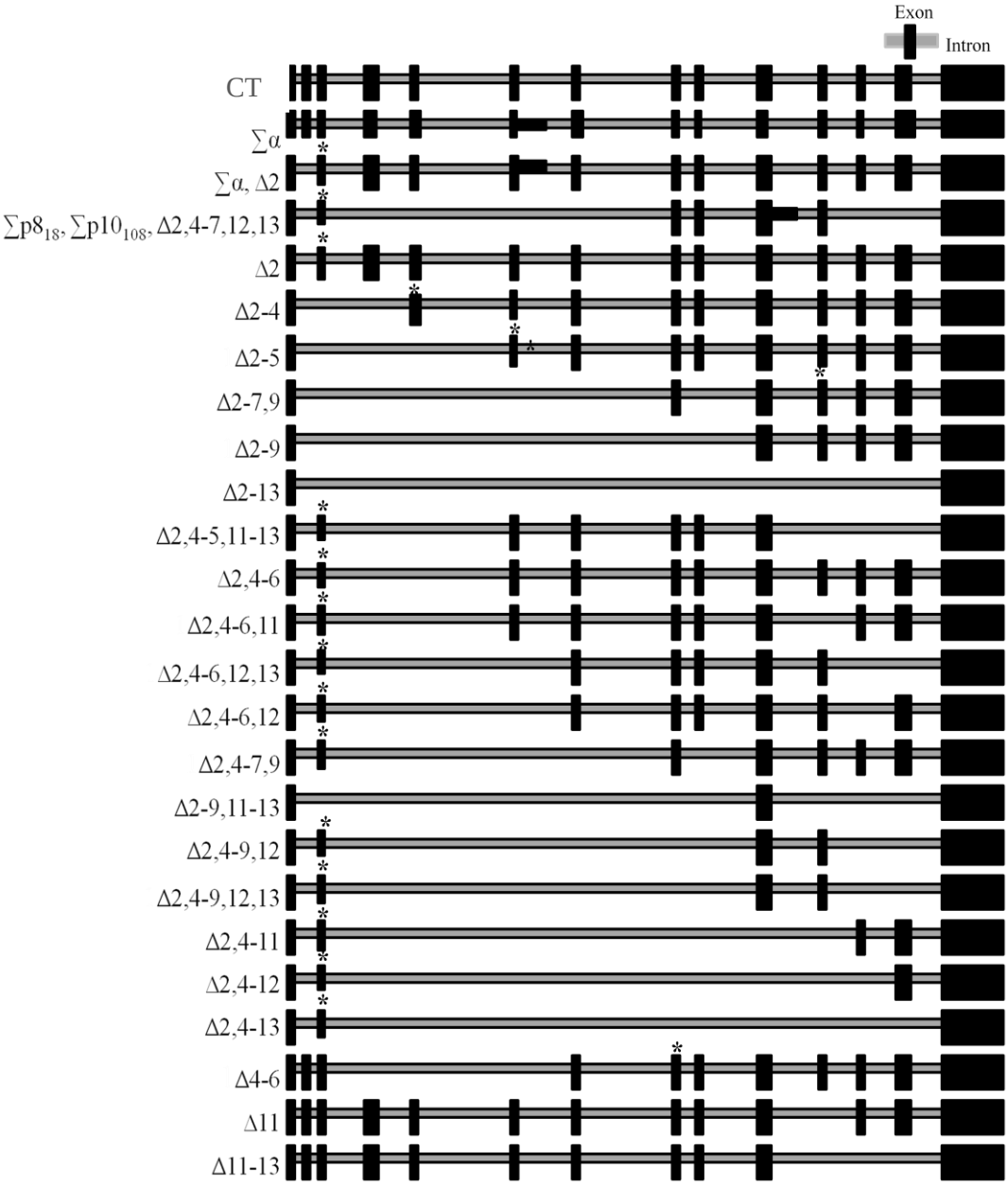

Figure 3.8 *POLB* transcript types characterized from the human brain. * Indicates location of a PTC.

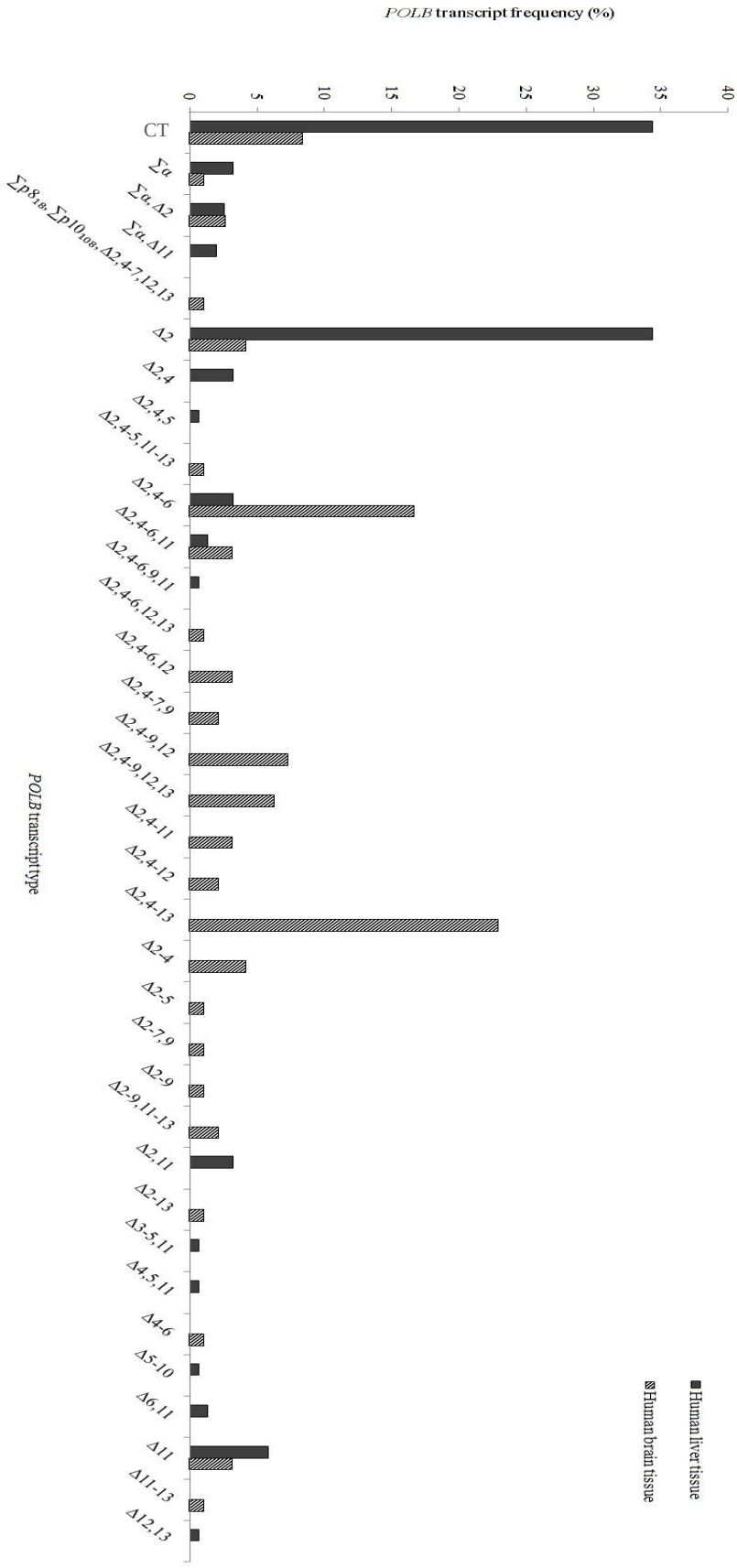

Figure 3.9 Frequency of *POLB* transcripts types in human liver and brain.

Table 3.2 *POLB* frequency of AS and transcript diversity in human liver and brain.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	53	65.4	18
	Splice variant	100		
Brain	Canonical transcript	8	91.7	25
	Splice variant	88		

AS of *TRPV1* in human tissues

TRPV1 cloned transcripts were successfully amplified from human liver and brain (Figures 3.10, lane 10 and 3.11, lane 2). The resulting size of the PCR product was approximately 2048 bps (Table 1.2). A total of 61 and 88 *TRPV1* transcripts were characterized from the human liver and brain, respectively.

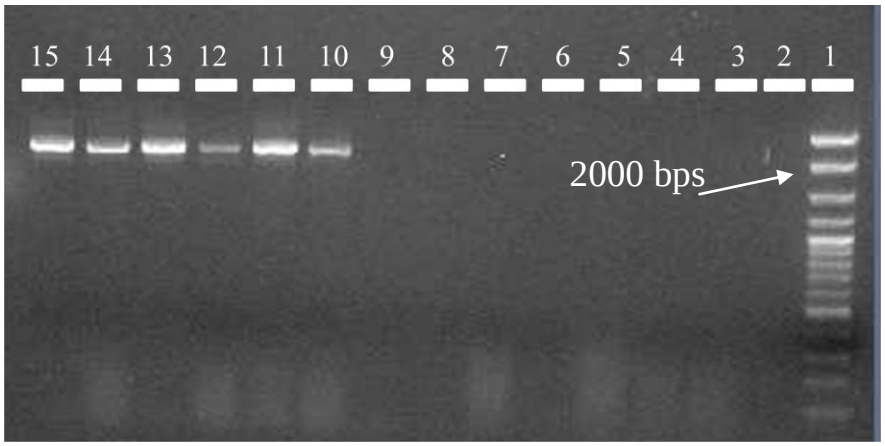

Figure 3.10 *TRPV1* PCR products amplified from human liver. Lane 1 shows the DNA size marker. Lane 2 contains a negative control where water was used in place of template. Lanes 3-9 show unsuccessful amplification of TRPV1 from the human liver using alternative *TRPV1* primers. Lanes 10, 11 and 14 show *TRPV1* PCR product from the human liver. Lanes 12, 13, and 15 show *TRPV1* PCR product from the human untransformed fibroblast cell line, MRC5.

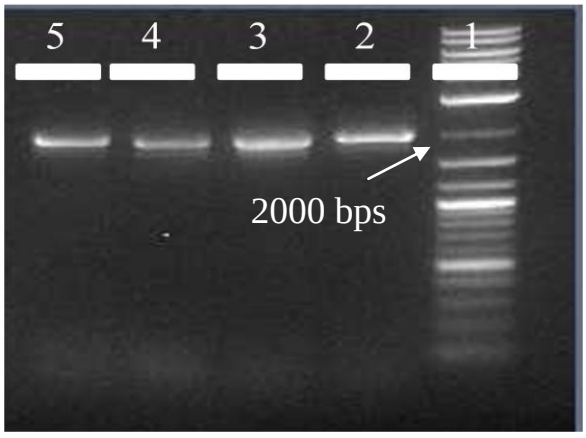

Figure 3.11 *TRPV1* PCR products amplified from human brain. Lane 1 shows the DNA size marker. Lanes 2-5 show *TRPV1* PCR product from the human brain.

TRPV1 AS in human liver tissue

Of the 61 *TRPV1* transcripts characterized from the human liver five were SVs. The *TRPV1* SV frequency in the human liver was found to be 8.19% (Table 3.3). The *TRPV1* transcript diversity of the human liver was three (Table 3.3).

The two *TRPV1* SVs observed in the human liver tissue were $\Sigma pI5_{47}$ and $\Delta 8$ (Figure 3.13). The $\Sigma pI5_{47}$ *TRPV1* SV contained a partial 47 bp retention of intron five and occurred at a frequency of 4.9% (3/61) (Figure 3.12). This 47 bp inclusion resulted in a PTC in exon six. The *TRPV1* $\Delta 8$ SV was lacking 180 bps compared to the canonical human *TRPV1* transcript. This SV occurred at a frequency of 3.3% (2/61) in the human liver (Figure 3.12). No PTC was created from the skip of exon eight.

TRPV1 AS in human brain tissue

Of the 88 *TRPV1* transcripts characterized from the human brain four were SVs. The *TRPV1* SV frequency in the human brain was 4.9% (3/61) (Table 3.3). The *TRPV1* transcript diversity of the human brain was four (Table 3.3). Two of the three *TRPV1* SVs observed in the human brain were previously described in the human liver. These SVs were $\Sigma 5_{47}$ and $\Delta 8$ which occurred at frequencies of 1.6% (1/61) and 2.3% (2/61), respectively in the human brain (Figures 3.12 and 3.13).

One *TRPV1* SV was found only in the human brain, $\Delta p8_{168}$,9,10,11,$p12_{77}$, which occurred at a frequency of 1.6% (1/61) (Figures 3.12 and 3.13) . Due to the multiple partial and multiple full exon skipping events in this SV a PTC was created in exon 13. There was a statistically significant difference in *TRPV1* SV frequencies between human liver and brain (x^2= 0.8463, d.f=1, p=0.3576,α=0.05).

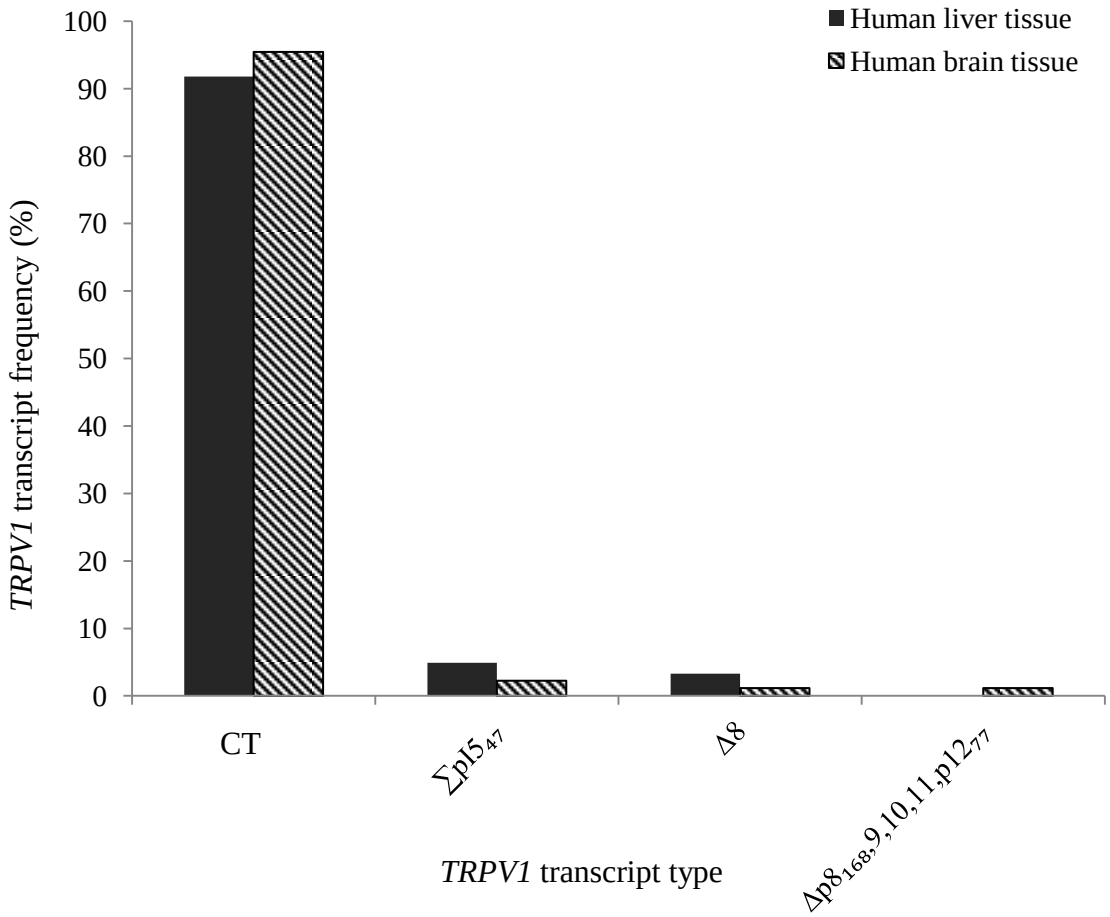

Figure 3.12 *TRPV1* CT and SV frequencies of the human liver and brain.

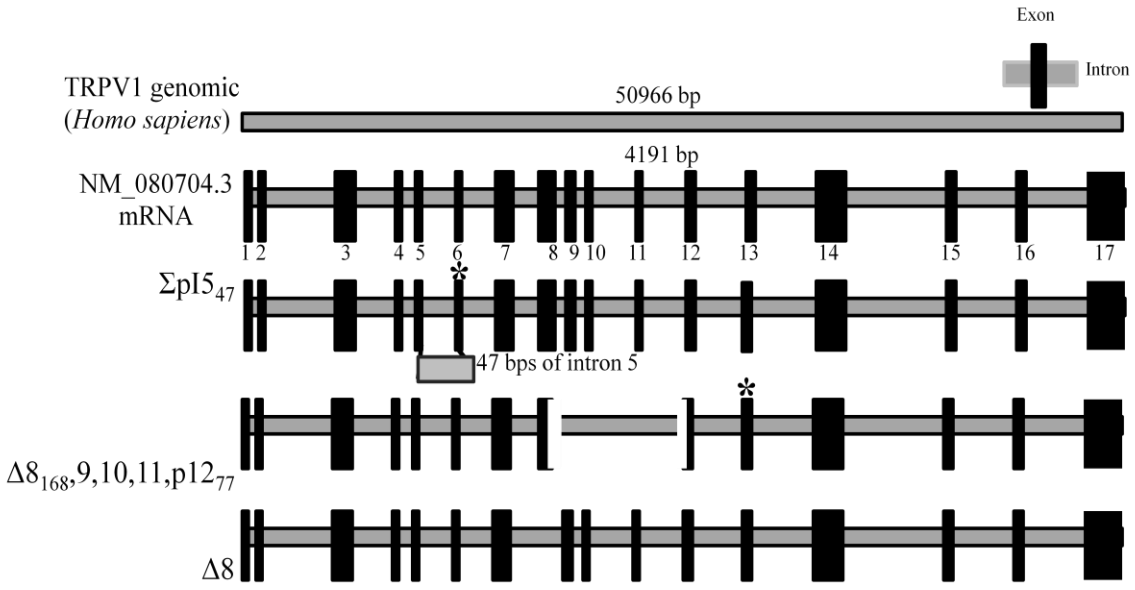

Figure 3.13 *TRPV1* transcript types characterized from the liver and brain of the human.

Table 3.3 *TRPV1* frequency of AS and transcript diversity in human liver and brain.

Tissue type	Type of transcript	Number of transcripts found	Splice variant frequency (%)	Diversity in tissue type
Liver	Canonical transcript	56	8.19	3
	Splice variant	5		
Brain	Canonical transcript	84	4.76	4
	Splice variant	4		

***POLB* and *TRPV1* SV frequencies in the liver and brain of the human**

There was a statistically significant difference found between *POLB* SV frequencies in human liver compared to human brain (x^2: 57.0224, d.f=1, p=<0.0001, α=0.05). *TRPV1* also showed a statistically significant difference between tissues types in both *TRPV1* SV frequency and transcript diversity (x^2: 139.3939, d.f=1, p=<0.0001, α=0.05).

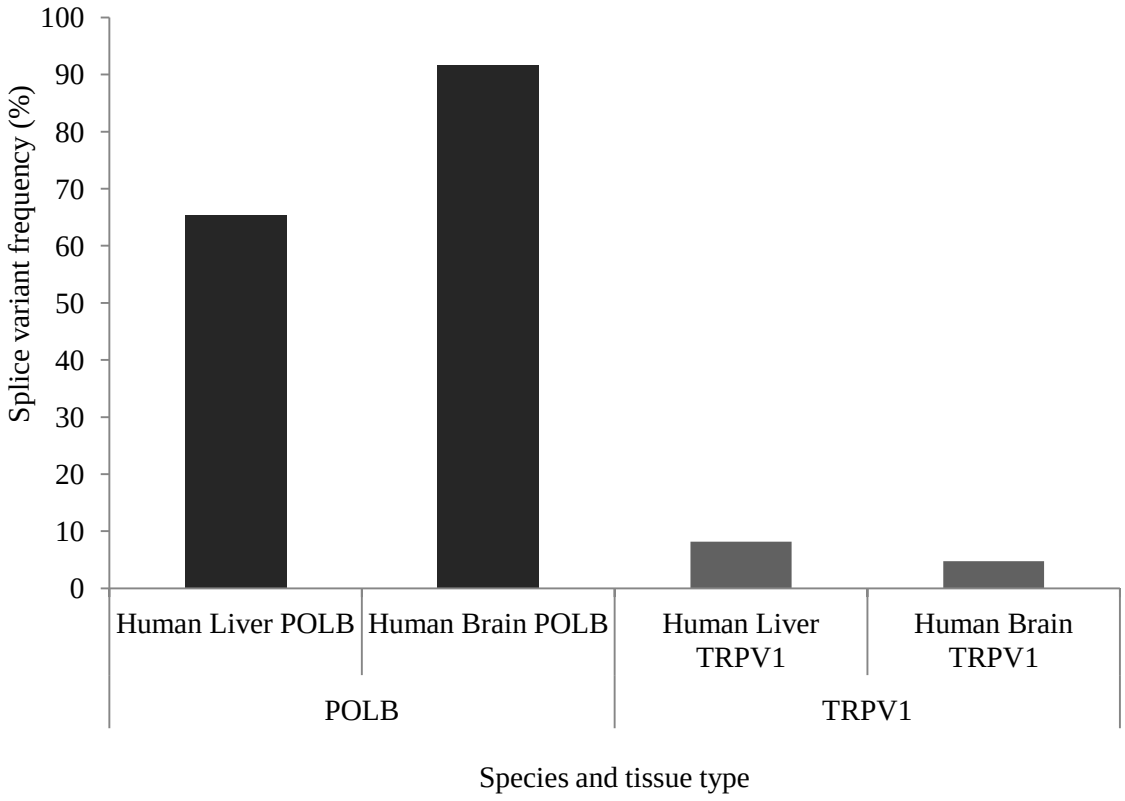

Figure 3.14 SV frequencies of *POLB* and *TRPV1* in the human liver and brain.

POLB AS in human tissues as compared to human cell lines

In human untransformed fibroblast cell lines *POLB* SV frequency was reported to be 57.4% with a transcript diversity of 14 (Skandalis et al., 2010). Comparing this data to the human liver tissue there was no statistically significant difference found in *POLB* SV frequencies between cell lines and liver tissue (x^2=2.63, d.f=1, p=0.10, α=0.05) (Figure 3.15). The human brain showed higher *POLB* SV frequencies than both human fibroblast cell lines and liver tissue (Figure 3.15).

POLB transcript diversity was lowest in cell lines with 14 different *POLB* transcripts types reported, compared to the 18 and 25 types observed in human liver and brain tissues, respectively (Figure 3.16).

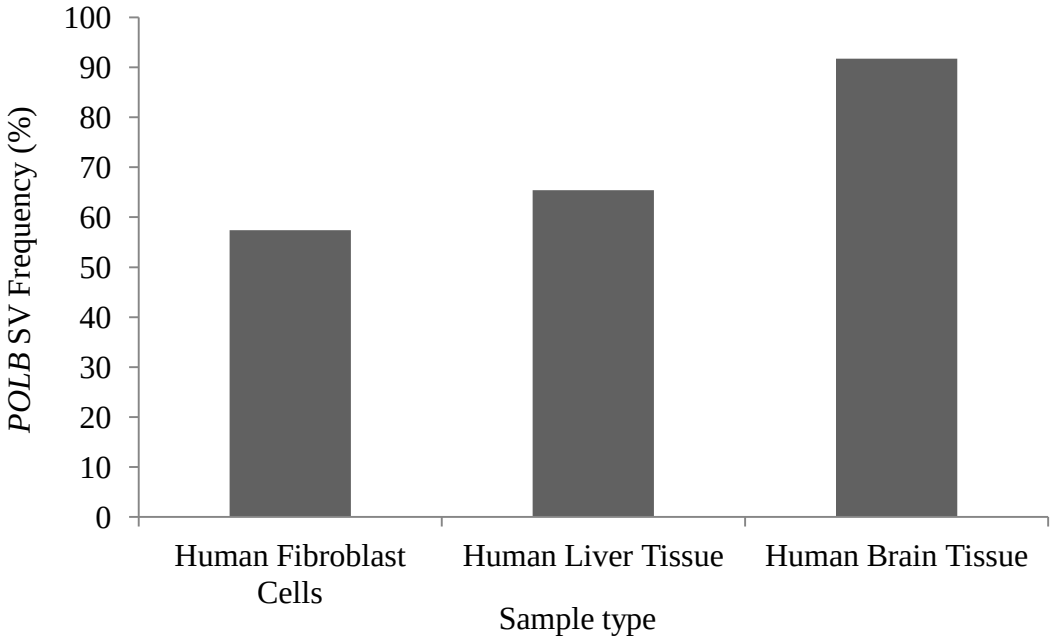

Figure 3.15 *POLB* SV frequencies and transcript diversities in human untransformed fibroblast cell lines, human liver and brain tissues.

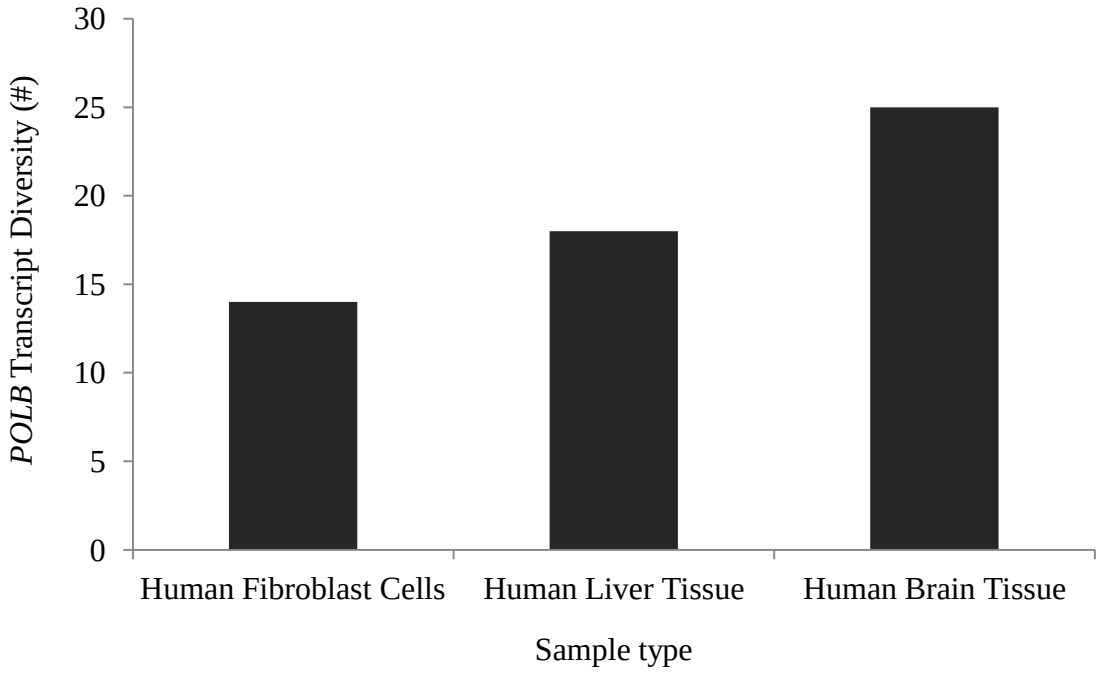

Figure 3.16 *POLB* transcript diversities in human untransformed fibroblast cell lines, human liver and brain tissues.

AS of *POLB* in non-mammalian and mammalian species tissues

Non-mammalian tissues showed low AS of *POLB* compared to mammalian tissues in all tissue types. All non-mammalian species investigated here exhibited low *POLB* SV frequencies, ≤3.2% compared to mammalian species which showed moderate to high SV frequencies (19.7 - 91.7%).

***POLB* AS and lifespan in mammals**

Taken together, all known mammalian data including primate cell lines and gathered from rabbit and human brain tissue, *POLB* SV frequency and transcript diversity correlate strongly with log lifespan (frequency: R^2=0.9223; diversity: R^2=0.7031) (Figure 3.17).

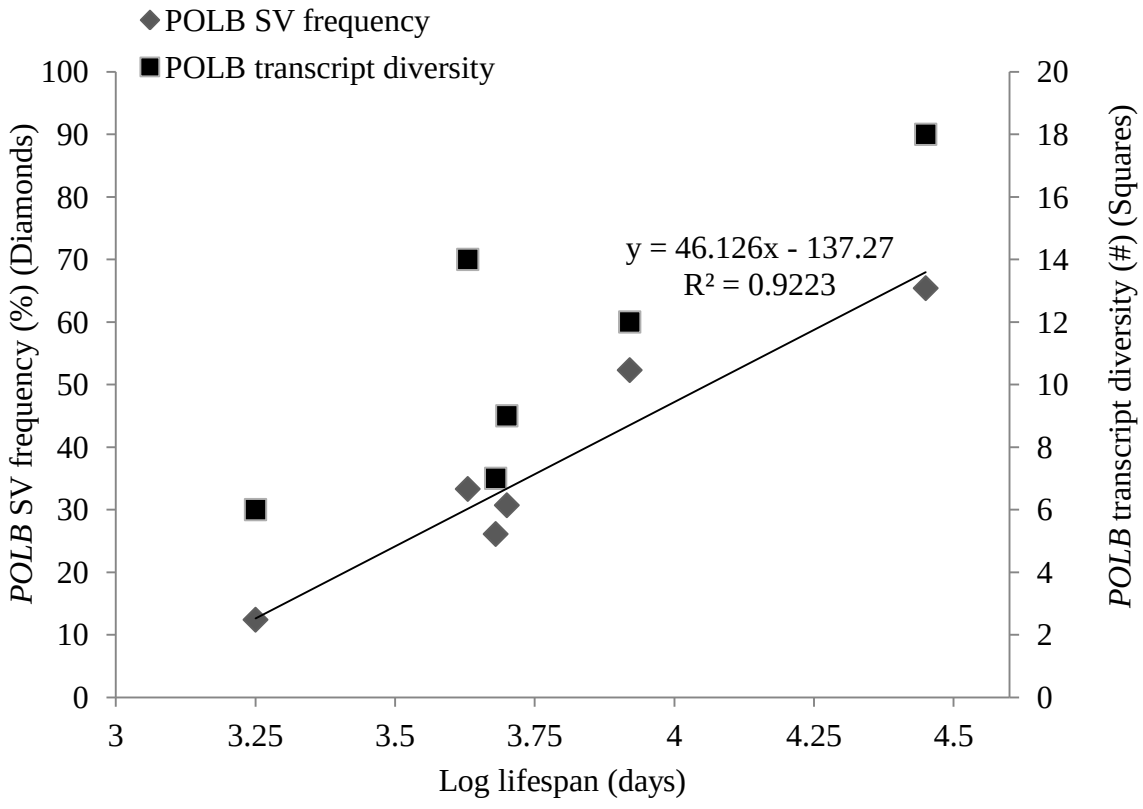

Figure 3.17 Mammalian *POLB* SV frequencies, transcript diversity and log lifespan. Human and rabbit *POLB* SV frequency and transcript diversity data collected from brain tissue.

DISCUSSION

Analysis and comparison of the AS patterns of *POLB* in human tissues to human fibroblast cell lines revealed that liver tissues and cell lines showed similar *POLB* SV frequency, whereas brain tissue did not. *POLB* SV frequency and transcript diversity showed statistically significant differences between human liver and brain tissues. Moreover, the human brain tissue had the highest *POLB* SV frequency and transcript diversity of all vertebrate tissues, at 91.7% and 25, respectively. Analysis of the AS patterns of the comparison gene, *TRPV1* to *POLB* patterns indicated that there was no tissue specific overall increase or decrease of AS in either human liver or brain tissue. Finally, a positive correlation between *POLB* SV frequency and lifespan was observed in mammals but not in all vertebrates.

AS patterns of *POLB* in rabbit liver and brain tissues

Rabbit liver and brain demonstrated similar *POLB* SV frequency, at 19.7% and 20%, respectively, with no statically significant difference found between the two tissues. *POLB* transcript diversity was equal in these tissues with eight unique transcripts. Similar to non-mammals no statistically significant tissue specific *POLB* AS was found in the rabbit. These data, in conjunction with the *POLB* AS patterns in non-mammals suggests the AS of *POLB* is not tissue specific with regards to liver or brain tissues, for the majority of vertebrate species.

The common *POLB* AS events in the rabbit tissues were exon two or exon four skip in combination with additional exon skip. For example, exon two was often skipped in combination with exons 9, 11 or 12. Alternatively, exon four was skipped along with exons 5,6,7,8,9 and/or 10. It is interesting to note that the multiple exon skips involved with exon two only occurred in the rabbit brain whereas exon four skip events were seen

in both tissues. This is similar to non-mammal exon two skip events which were primarily observed in brain tissue. Exon two skip was the only *POLB* splicing event which appeared to be conserved throughout vertebrate species examined. The exon 11 SV was also recorded in rabbit tissues. This SV has been previously reported in human tissues and cell lines but this represents the first evidence of the event in non-mammals (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002).

AS patterns of *POLB* in human liver and brain tissues

In human tissues *POLB* was alternatively spliced in a tissue specific manner. High SV frequencies were observed in the human brain compared to the liver, 91.7% versus 65.4%, respectively. Transcript diversity was also higher in the human brain at 25 types compared to the liver which had 18. These tissues have the highest recorded *POLB* SV frequencies and diversities of all vertebrate species. Six *POLB* SVs were common to both liver and brain tissues in humans. These SVs were Σα, Σα,Δ2, Δ2, Δ2,4-6, Δ2,4-6,11 and Δ11. All of these SVs have been observed human cell lines previously (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002).

***POLB* AS patterns in cell lines and tissues**

The first hypothesis of this chapter was that *POLB* AS data collected from human fibroblast cell lines were consistent with RNA isolated from human tissues. The observed results do not agree with this hypothesis. Although the *POLB* SV frequencies between the human liver and fibroblast cell lines are similar the types observed are different. A large difference between *POLB* AS in the brain compared to cell lines was also observed, 91.7% versus 57.4%. There were 17 novel *POLB* SVs characterized in human tissues

which had not been observed in human cell lines or tissues on previous occasions (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002).

There are several possibilities as to why these novel *POLB* SVs were present in tissues and yet not found in cell lines. First, cell type may have influenced the types of *POLB* SVs observed. If *POLB* AS is cell type specific perhaps different cell types produce specific SVs. Indeed, AS patterns of numerous genes have been shown to be cell type specific (Hanamura et al., 1998). Specifically, *POLB* has been shown to produce specific SV types between untransformed normal foetal lung fibroblast cells and transformed human lymphoblastoid cells (Disher et al., 2007). The results shown here demonstrate AS patterns of *POLB* do differ between cell lines and tissues. Moreover, tissue specific splicing patterns of *POLB* were observed in human tissues, which would not be represented in a single cell type analysis. These data indicate that cell lines do not represent the *POLB* splicing patterns in all human tissues.

The Δ2 *POLB* SV in vertebrate tissues

The human liver showed equal frequencies of CT *POLB* and Δ2. If the SV frequencies found reflect the expression levels it could suggest that 50% of the *POLB* transcripts expressed in the liver are Δ2. The Δ2 variant is by far the most studied *POLB* SV. The Δ2 variant has been reported in five primate species, including humans as well as numerous human cell lines (Skandalis et al., 2010; Simonelli et al., 2009; Disher et al., 2007; Skandalis et al., 2004; Thompson et al., 2002). This thesis is the first work to show that the Δ2 *POLB* SV is primarily found in brain tissue and not liver tissue for the

majority of vertebrate species, with the exception of humans. Human brain tissue showed Δ2 occurred at a frequency of 4.17%.

Simonelli et al., (2009) investigated Δ2 frequencies in various human tissues and suggest the skip of exon two observed was not tissue specific or disease related. The findings presented in this thesis suggest that exon two skip event is tissue specific, at least in the liver and brain tissues of non-mammals, gull, frog and bearded dragon and one mammal (rabbit). Whether the Δ2 variant is present in other tissues of non-mammals such as heart or muscle tissues is not known at this time. The question remains as to why it was observed in brain tissues and not liver in these species.

One possibility may be that the Δ2 SV plays a key role in the down regulation of normal *POLB* activity in the vertebrate brain. Studies have demonstrated that several tumour types over express *POLB* (reviewed by Starcevic et al., 2004), and that high levels of *POLB* destabilize some cellular processes that lead to tumour progression (Sweasy et al., 2003). Perhaps the expression of Δ2 functions to regulate the expression levels of *POLB* in the brain to counter tumourgenesis. Srivastava et al., (1998) showed that in mammals POLB protein is the rate limiting step of BER *in vitro*. Results in this chapter showed that Δ2 and CT *POLB* were present in the human brain. Based on this finding and considering that POLB is the rate limiting step in BER in mammals we may expect to find reduced BER in the human brain if Δ2 is acting to down regulate normal POLB expression. Lu et al. (2004) suggested that BER is in fact reduced in the human brain during aging, specifically in individuals over the age of 40. All of the donors of the pooled human brain RNA used in the work presented here were over the age of 40.

POLB frequency and transcript diversity was highest in human brain tissue

Considering the ages of donors of the human RNA the reason why *POLB* SVs levels exist in the brain could simply be a result of advanced age. The AS of certain genes has been shown to increase with age in humans (Tollervey et al., 2011). A future investigation examining the SV frequencies, types and expression levels of *POLB* SVs in individuals of different ages in multiple tissue types would be a very informative.

Another possibility for high *POLB* SV frequency and transcript diversity in the human brain is the circumstances of the RNA isolation. Awasthi et al. (1997) observed that oxidative stress significantly increased in rat brains within 60 minutes following brain injury. Similar increases in oxidative stress following brain injury have been reported in humans (Clausen et al., 2011). If trauma occurred to the brain prior to RNA isolation the effects of increased oxidative stress may have influenced AS patterns of genes in this tissue. Oxidative stress has been demonstrated to affect the AS of *POLB* (Disher et al., 2007). Disher et al. (2007) observed that under oxidative stress the overall frequency of *POLB* SVs increased in both untransformed normal foetal lung fibroblast cells and transformed human lymphoblastoid cells. It could be possible that the high *POLB* SV frequency and transcript diversity observed in the human brain was a result of an increase in oxidative stress in donor brain tissue.

Finally, *POLB* SV frequencies may have been higher in human brain tissues as compared to liver tissue because the human brain has a higher overall SV frequency. This higher SV frequency could apply to all genes in the brain or perhaps just DNA repair genes. In humans, out of several tissue types, liver and brain tissues showed the highest levels of AS at 42% and 39%, respectively (Yeo et al., 2004). This analysis was based on

the number of genes which showed one or more AS event. Although Yeo et al. (2004) indicated that the number of genes which undergo AS was highest in the human brain, they did not report the SV frequencies. If SV frequencies are higher for all genes in the brain, additional genes besides *POLB* should show high SV frequencies when compared to liver tissues. To investigate this in human liver and brain tissues *POLB* SV frequencies were compared to a second gene, *TRPV1*. Results showed that *TRPV1* SV frequency and transcript diversity were higher in the liver compared to the brain suggesting the brain does not have overall higher AS frequencies with respect to *POLB* and *TRPV1*. It is possible that DNA repair genes have high AS frequencies in the human brain as compared to non-DNA repair genes like *TRPV1*. Future experiments could investigate the AS of a variety of DNA repair genes in different human tissues to see whether all DNA repair genes have high frequencies of AS or just *POLB*.

AS of *POLB, TRPV1* and a possible link through inflammation

It is possible that *TRPV1* generally has a higher SV frequency in the liver tissue compared to the brain. Chicken liver tissue showed a very high level of *TRPV1* splicing compared to the brain however; frog tissues demonstrated the opposite trend. Taken together, these data suggest that *POLB* SV frequency and transcript diversity is high in the human brain compared to the human liver and it was not due to an overall systemic increase of AS in the tissue. Interestingly, two novel *TRPV1* SVs out of the three SVs were observed in human tissues. The Δ8 *TRPV1* SV, described by Lu et al. (2005) was named hTRPV1b. The hTRPV1b SV was shown to produce a functional protein which is activated by temperature but not capsaicin or protons (Lu et al., 2005). Unlike hTRPV1b

the two novel human *TRPV1* SVs, $\Delta p8_{168},9,10,11,p12_{77}$ and $\Sigma 5_{47}$ both harbour PTCs in their transcripts. Whether these SVs have any obvious function is not known at this time.

TRPV1 was chosen as a comparison gene to *POLB* in this thesis based on considerations discussed in Chapter Two. However, there is a possible link between these two genes that is noteworthy. When tissue damage occurs signalling pathways are induced that trigger inflammation (Durnett, et al., 2002). In turn, an increase in inflammatory agents result in the direct activation of TRPV1 channels which cause TRPV1-mediated [Ca^{2+}] to rise triggering a release of pro-inflammatory agents (Planells-Cases et al., 2005). Inflammation induces oxidative stress by generating reactive oxygen species (ROS) (Bartsch et al., 2006; reviewed by Halliwell et al., 1996). DNA damage caused by ROS is well documented (reviewed by Cooke et al., 2003). The overall frequency of *POLB* SVs is increased in cells exposed to oxidative stress (Disher et al., 2007). Given the inflammation and oxidative stress link between *POLB* and *TRPV1* it would be intriguing to investigate the splicing patterns of the two genes during exposure to inflammation.

POLB transcript diversity in mammals

POLB transcript diversity was higher in mammalian as compared to non-mammalian tissues. There were 13 different *POLB* SVs types observed in the rabbit tissues and 34 types in the human tissues. The majority of human brain *POLB* SVs had exon two skip in combination with multiple other exon skips. Several of these *POLB* SVs were characterized with large portions of the gene skipped. Thompson et al. (2002) described *POLB* SVs with multiple exon skips, in some cases 12 out of 14 exons, in bladder tumour tissues. Similar *POLB* SVs were observed in gastric cancer cell lines by

Simonelli et al. (2009). Why *POLB* displays these splicing patterns in this tissue is unknown. The presence of multiple exon skipping in mammalian verses non-mammalian *POLB* SVs agrees with previous findings suggesting that the occurrence of exon skipping increases further up the eukaryotic evolutionary tree (reviewed by Keren et al., 2010 and Kim et al., 2007).

Premature termination codons (PTCs) were found in all but one *POLB* SV in non-mammals, with chicken liver tissues demonstrating the only *POLB* SVs without a PTC. In rabbit and human tissues 66.6% and 66.4% of the *POLB* SVs generated contained PTCs, respectively. PTCs have been observed in a considerable fraction of generated SVs and an estimated 12% of human genes which undergo AS have them (reviewed by Lareau et al., 2007). In some cases splicing errors, such as mutations in introns or incorrect recognition of splice sites by the splicing machinery, could result in the creation of a PTC in SV (Lareau et al., 2007). Alternatively, an estimated 66% of human genes which generate SVs containing a PTC are a result of a PTC specific splice site (Soergel 2005). It has been suggested that SVs which contain a PTC could be functional and perhaps be conserved in a species specific manner (Kan et al., 2002). Analyses by Skandalis et al. (2004) revealed that at least a portion of *POLB* SVs generated are a result of aberrant splicing. The evolutionary conservation of the Δ2 variant suggests that some of the *POLB* SVs containing PTCs are not subject to NMD. Moreover, the *POLB* SVs containing PTCs are perhaps selected by the cell to provide an additional level of regulation, for example as a micro-RNA precursor (Simonelli et al., 2009).

A positive correlation between *POLB* SV frequency and lifespan in mammals

Mammalian tissue data support the hypothesis that there is a positive correlation between *POLB* SV frequency and lifespan in mammals. This hypothesis agrees with the findings reported in primate species and mammals in general as suggested by the Skandalis et al. (2010) (unpublished). These data suggest that as lifespan of a species increases the *POLB* SV frequency and transcript diversity does as well. The mechanism behind this correlation has been proposed to involve down-regulation of gene expression (Skandalis et al., 2010) so perhaps this results in a reduction in CT *POLB* synthesis.

Interestingly, Brown et al. (2007) observed a negative correlation between POLB protein activity and species maximum lifespan in mammalian primary dermal fibroblasts under atmospheric oxygen. Although the reported increase in *POLB* SVs and a decrease in POLB activity with lifespan of a species (Brown et al., 2007) cannot be definitively linked to the AS of *POLB* without expression levels of *POLB* transcripts, (both CT and SVs) these results may indicate a general trend is present in mammals. Additionally, oxygen conditions of the fibroblast cell lines were shown to influence POLB activity (Brown et al., 2007) which further demonstrates the importance of cell line and tissue comparison experiments. Furthermore, it was proposed by Page et al. (2011) that the relationships between BER and lifespan may be cell type specific.

Particular *POLB* SVs are suggested to act as dominant negative regulators of normal POLB activity (Thompson et al., 2002) and it was demonstrated that high levels of POLB destabilize further cellular processes and in turn lead to tumour progression (Sweasy et al., 2003). Perhaps some *POLB* SVs, such as Δ2 and Δ11, have functional roles without forming functional protein but, we must ask, what about the remainder of

the SVs? Moreover, why would AS of *POLB* correlate to a longer lifespan when *POLB* functions as a DNA repair protein? Perhaps the process of generating *POLB* SVs alone is related to lifespan and not the production of specific SVs themselves, as suggested by Skandalis et al. (2010). With analysis of only two mammalian species it is difficult to determine whether particular *POLB* SVs are evolutionary conserved in mammals, apart from Δ2. However, in addition to Δ2 four other *POLB* SVs were observed in both rabbit and human tissues, Δ2,11, Δ4-6, Δ11 and Δ12,13. Future experiments could focus on an investigation of these four SVs in other mammalian tissues. The conservation of these SVs could suggest a functional role for them in mammals.

The final proposed hypothesis was that a positive correlation existed between the AS of *POLB* and lifespan in vertebrate species in general. Data gathered from four non-mammalian and two mammalian species taken together with observations in primate cell lines did not support this hypothesis. Instead, these data suggest that a positive correlation exist in mammals and not in non-mammals.

CONCLUSIONS

The AS patterns of *POLB* in mammals showed that liver and brain tissues had higher SV frequency and transcript diversity compared to non-mammalian tissues. The AS of *POLB* was found to be tissue specific in human liver and brain whereas in rabbit tissues it was not. There were 17 novel *POLB* SVs in human tissues which had not been described previously. Additionally, all the *POLB* SVs characterized in rabbit tissues have not been reported prior to this work. The human brain demonstrated the highest *POLB* SV frequency and transcript diversity in all vertebrate species investigated. Investigation of the AS patterns of a second gene *TRPV1*, suggests that the *POLB* AS patterns observed are not a result of an overall increase in AS in the liver or brain tissues of mammals. The positive correlation between the AS of *POLB* and lifespan previously reported in primates appears to exist in all mammals. However, taken together with non-mammalian *POLB* AS patterns a positive correlation does not appear to exist in all vertebrates.

General conclusions and further directions

The main objective of this work was to investigate AS of the DNA polymerase gene, *POLB*, in vertebrates. Previous work demonstrates that a positive correlation between *POLB* SV frequency and lifespan exists in primate cell lines (Skandalis et al., 2010). Whether this positive correlation exists in all vertebrates is unknown. This work aimed to explore said correlation. Does this positive correlation exist in non-mammals? Do human cell lines reflect the AS patterns of *POLB* seen in human tissues? Is the AS of *POLB* representative of the SV frequencies of other genes such as non-DNA repair genes?

To answer these questions, four non-mammalian species were selected: the chicken, gull, frog and bearded dragon, as well as two mammalian species: the rabbit and human. *POLB* transcripts were characterized in liver and brain tissues from aforementioned vertebrate species to determine *POLB* SV frequency and transcript diversity.

The second chapter of this thesis revealed that AS of *POLB* in non-mammalian tissues was low, both in frequency and transcript diversity, compared to observations made in primate cell lines. In terms of liver and brain tissues, statistically significant differences in *POLB* SV frequency were found in the bearded dragon and gull but not in chicken and frog. Analysis of the comparison gene, *TRPV1*, suggested that *POLB* was not subject to an overall increase or decrease of AS within particular tissues. Comparing *POLB* SV frequency with lifespan revealed that a positive correlation between the AS of *POLB* and lifespan did not exist in non-mammalian tissues. Finally, the skip of exon two in various forms was observed in three out of four species but only in brain tissues.

The third chapter of this thesis investigated the AS of *POLB* in mammals and cell lines. *POLB* SV frequency was similar in human liver tissue compared with cell lines, while transcript diversities were not. The human brain showed significantly higher *POLB* SV frequency and transcript diversity than cell lines. While it appears that human untransformed fibroblast cell lines may be an appropriate representation for some human tissues, they show an underestimation of *POLB* AS for others. *TRPV1* analysis suggested that there was no overall increase or decrease in AS in liver or brain tissues of humans, which would have affected the AS patterns of *POLB*. Tissue specific AS of *POLB* seemed to exist in human tissues and not in rabbit tissues.

Finally, the data in chapters two and three suggest that a positive correlation between the AS of *POLB* and lifespan exists in mammals but not all vertebrates. Data from the investigation of POLB protein activity in mammals showed a decrease in POLB activity in mammalian species with longer lifespan. This may suggest that the decrease in POLB DNA repair capacity results from the increase in *POLB* SVs in mammals (Brown et al., 2007). Furthermore, Page et al. (2011) did not find that the increase in POLB activity in brain or liver tissues of vertebrate species evolved with lifespan. Perhaps the decrease in POLB activity is a result of increased unproductive *POLB* SV production. Such a relationship could be connected to lifespan in mammals, but not non-mammals, and therefore not in vertebrates as a whole. To investigate this relationship, expression levels of *POLB* CT and SV transcripts could be collected and examined to elucidate the connection between the production of SVs and protein expression.

There are several avenues of research to investigate that could be investigated to continue the work presented in this thesis. For example, although *POLB* is evolutionary

conserved at the nucleotide level, mutations may be present in splice sites of non-mammals as compared to mammals. If there are changes in the splice sites this presents a possible explanation for the low *POLB* SV frequency observed in non-mammals. Bioinformatic analyses of the splice sites of *POLB* from numerous vertebrates would be the first step to elucidate this possibility.

In terms of tissue collection, each species was represented by a single organism with the exception of the human brain. In the case of the human brain, RNA was isolated and pooled from 12 individuals. Expanding the investigation to more individuals of each species should reveal variations within *POLB* SV between individuals. Furthermore, RNA isolation from distinct human brain tissue samples should be collected. This would allow data from these samples to be compared to human pooled brain RNA data to help elucidate possible variations in AS patterns of *POLB* present between individuals.

Investigating the expression levels of *POLB* SVs in tissues would be incredibly useful for understanding the AS of *POLB*. The presence and frequencies of *POLB* SVs alone are not enough to indicate their expression levels. Using quantitative PCR to measure the expression levels of each SV would indicate whether *POLB* SVs are significant enough to influence the overall DNA repair capacity of normal *POLB*. Despite the considerable time and effort required to accomplish such a task, results would provide valuable insight into the AS of *POLB* in vertebrates.

References

Abbotts, J., SenGupta, D. N., Zmudka, B., Widen, S. G., Notario, V., Wilson, S. H. 1988. Expression of human DNA polymerase. beta. in Escherichia coli and characterization of the recombinant enzyme. Biochemistry 27: 901-909.

Alló, M. 2009. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nature Struct. Mol. Biol. 16: 717–724.

Artamonova II, Gelfand M. S. 2007. Comparative genomics and evolution of alternative splicing: The pessimists' science. Chem Rev 107: 3407–3430.

Awasthi, D., Church, D. F., Torbati, D., Carey, M. E., Pryor, W. A. 1997.Oxidative stress following traumatic brain injury in rats. Surg. Neurol.47: 575–581.

Bannwarth, S., and A. Gatignol. 2005. HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. Curr. HIV Res. 3:61-71.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield for λ bacteriophage templates. Proc. Natl. Acad. Sci. USA 91: 2216–2220.

Bell T. J., Miyashiro K. Y., Sul J. Y., Buckley P. T., Lee M. T., McCullough R., Jochems J., Kim J., Cantor C. R., Parsons T. D., Eberwine J. H. 2010. Intron retention facilitates splice variant diversity in calcium-activated big potassium channel populations. Proc Natl Acad Sci USA 107:21152-21157.

Bhattacharyya N., Banerjee, S. 1997. A variant of DNA polymerase β acts as a dominant negative mutant. Proc. Natl Acad. Sci. USA 94: 10324–10329.

Bhattacharyya, N., Banerjee, S. 2001. A novel role of XRCC1 in the functions of a DNA polymerase beta variant. Biochemistry 40: 9005-9013.

Bhattacharyya, N., Chen, H. C., Grundfest-Broniatowski, S. & Banerjee, S. 1999. Variant forms of DNA polymerase beta in primary lung carcinomas. Biochem. Biophys. Res. Commun. 259: 429-435.

Black, D. 2003. Mechanisms of alternative pre-messenger RNA splicing. Annual Review of Biochemistry 72: 291-336.

Bohlen C.J., Priel A., Zhou S., King D., Siemens J, Julius D. 2010. A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain Cell 141: 834-845.

Boue S, Letunic I, Bork P. 2003. Alternative splicing and evolution. Bioessays 25: 1031-1034.

Bowen R. L., Atwood C. S. 2004. Living and dying for sex: a theory of aging based on the modulation of cell cycle signaling by reproductive hormones. Gerontology 50: 265-90.

Brett D., Pospisil,H., Valcarcel,J., Reich,J. and Bork,P. 2002. Alternative splicing and genome complexity. Nature Genet., 30: 29–30.

Clausen F., Marklund N., Lewén A., Enblad P., Basu S., Hillered L. 2011. Interstitial F2-isoprostane 8-iso-PGF2α as a biomarker of oxidative stress following severe human traumatic brain injury. Published online.

Caterina, M. J., Julius D. 1997. The capsaicin receptor: a heat-activated ion channel in the pain pathway. Nature 389: 816–824.

Caterina, M. J., Leffler A., Malmberg A. B., Martin W. J., Trafton J., Petersen-Zeitz K. R., Koltzenburg M., Basbaum A. I, Julius D. 2000. Impaired nociception and pain sensation in mice lacking the capsaicin receptor. Science 288: 306–313.

Cavanaugh, D., Chesler A. T., Jackson A. C., Sigal Y., Yamanaka H., Grant R., O'Donnell R., Nicoll R. A., Shah. N., Julius D., Basbaum A. I. 2011. Trpv1 reporter mice reveal highly restricted brain distribution and functional expression in arteriolar smooth muscle. J. Neurosci. 31: 5067–5077.

Chan K, Houlbrook S, Zhang QM, Harrison M, Hickson ID, Dianov GL. 2007. Overexpression of DNA polymerase b results in an increased rate of frameshift mutations during base skip repair. Mutagenesis 22: 183–188.

Chen Y E, Drucker D J. 1997. Tissue-specific expression of unique mRNAs that encode proglucagon-derived peptides or exendin 4 in the lizard. J Biol Chem 272: 4108–15.

Chow, L. T.; Gelinas, R. E.; Broker, T. R.; Roberts, R. J. 1977. Cell: 12, 1-8.

Chyan Y. J., Strauss P. R., Wood T. G., Wilson S. H. 1996. Identification of novel mRNA isoforms for human DNA polymerase β. DNA Cell Biol 15: 653–659.

Cooke M. S., Evans M. D., Dizdaroglu M., Lunec J. 2003. Oxidative DNA damage: mechanisms, mutation, and disease FASEB J.17: 1195–1214.

Crick F.1970. Central dogma of molecular biology. Nature 227:561–563.

Cuccurese, M., Russo G., Russo A., Pietropaolo C. 2005. Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression. Nucleic Acids Res. 33:5965–5977.

Disher K, Skandalis A. 2007. Evidence of the modulation of mRNA splicing fidelity in humans by oxidative stress and p53. Genome 50: 946–953.

Durand C, Roeth R, Dweep H, Vlatkovic I, Decker E, Schneider KU. 2011. Alternative splicing and nonsensemediated RNA decay contribute to the regulation of SHOX expression. PLoS ONE 6:18115.

Durnett, Richardson J., Vasko M. R. 2002. Cellular mechanisms of neurogenic inflammation. J Pharmacol Exp Ther 304:88–98.

Eichner J., Zeller G., Laubinger S., Rätsch G. 2011. Support vector machines-based identification of alternative splicing in *Arabidopsis thaliana* from whole-genome tiling arrays. BMC Bioinformatics 12: 55

Eilers H., Lee S. Y., Hau C. W., Logvinova A., Schumacher M. A. 2007. The rat vanilloid receptor splice variant VR.5'sv blocks TRPV1 activiation. Neuroreport18: 969-73.

Evans T. C. 2008. New England Biolabs Summer. vol.3.2.

Fedoroff S., Richardson A. 1997. Protocols for Neural Cell Culture. Human Press. 131–141.

Filichkin S. A., Priest H. D., Givan S. A., Shen R., Bryant D. W., Fox S. E., Wong W.K., Mockler T. C. 2010. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. Genome Res 20: 45–58.

Florea L. 2006. Bioinformatics of alternative splicing and its regulation. Brief Bioinform 7: 55-69.

Gabut M, Samavarchi-Tehrani P, Wang X, Slobodeniuc V, O'Hanlon D, Sung HK, Alvarez M, Talukder S, Pan Q, Mazzoni E O. 2011. An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell 147:132-146.

Gaidatzis D., Jacobeit K., Oakeley E. J., Stadler M. B. 2009. Overestimation of alternative splicing caused by variable probe characteristics in exon arrays. Nucleic Acid Res. 37:107

Gavrilov, L. A., Gavrilova, N. S. 1992. The biology of life span: a quantitative approach. Harwood Acad. Publishers, New York.

Graveley BR. 2004. Mutually Exclusive Splicing of the Insect Dscam Pre-mRNA Directed by Competing Intronic RNA Secondary Structures. RNA 123: 65-73.

Gracheva E. O., Cordero-Morales J. F., González-Carcacía J. A, Ingolia T. A., Manno C., Aranguren C. I., Weissman J. S., Julius D. 2011. Ganglion-specific splicing of *TRPV1* underlies infrared sensation in vampire bats. Nature 476: 88-91.

Gutteridge J. Halliwell, B. 1996. Free Radicals and Antioxidants in the Year 2000. Free Radicals in Biology and Medicine 136-147

Hanamura A., Caceres J. F., Mayeda A., Franza B. R., Krainer A. R.1998. Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. RNA 4: 430–444.

Hawkins N. S., Hearn J., Evans H. R. 1991. Comparison of the capsaicin- and amino acid-sensitivity of dorsal root C fibres in the rat and the toad. Comp. Biochem. Physiol. 99: 513–516.

Hwang J., Maquat L. E. 2011. Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. 21: 422–430.

Idriss, H. T., Al-Assar, O., Wilson, S. H. 2002. DNA polymerase β. Int. J. Biochem. Cell Biol 34: 321–324.

Iida K., Go M. 2006. Survey of conserved alternative splicing events of mRNAs encoding SR proteins in land plants. Mol. Biol. Evol. 23:1085–1094.

Jordt S. E., Julius D. 2002. Molecular Basis for Species-Specific Sensitivity to "Hot" Chili Peppers. Cell 108: 421–430.

Kan Z., States D., Gish W. 2002. Selecting for functional alternative splices in ESTs. Genome Res 12:1837-1845.

Kemp, P.R., Ellis, P.D., Smith, C.W. 2005. Visualization of alternative splicing in vivo. Methods 37: 360–367.

Keren H., Lev-Maor G.,Ast G. 2010. Alternative splicing and evolution: diversification, exon definition and function. Nat. Rev. Genet. 11:345-355.

Kim, E., Magen, A., Ast, G. 2007. Different levels of alternative splicing among eukaryotes Nucleic Acids Res. 35: 125-131.

Kim, H., Klein, R., Majewski, J., and Ott, J. 2004. Estimating rates of alternative splicing in mammals and invertebrates. Nat. Genet. 36: 915–916.

Kim N., Alekseyenko A. V., Roy M., Lee C. 2007. The ASAP II database: Analysis and comparative genomics of alternative splicing in 15 animal species. Nucleic Acids Res. 2007: 93–98.

Krainer AR, Conway GC, Kozak D. 1990a. The essential pre-mRNA splicing factor SF2 influences 59 splice site selection by activating proximal sites. Cell 62:35 –42.

Lareau L. F., Inada M., Green R. E., Wengrod J. C., Brenner S. E. 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446: 926–929.

Lemaire M, Prime J, Ducommum B, Bugler B. 2004. Evolutionary conservation of a novel splice variant of the Cds1/CHK2 checkpoint kinase restricted to its regulatory domain. Cell Cycle 3:1267–70

Lewis BP, Green RE, Brenner SE. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A 100:189–92.

Li C., Kato M., Shiue L., Shively E. J., Ares Jr M., Lin R-J. 2006. Cell type and culture condition-dependent alternative splicing in human breast cancer cells revealed by splicing-sensitive microarrays. Cancer Res 66: 1990–1999.

Loeb L.A , 2001. A mutator phenotype in cancer. Cancer Res. 61: 3230.

Luco R. F.,Allo M.,Schor I. E., Kornblihtt A. R.,Misteli T. 2011. Epigenetics in alternative pre-mRNA splicing. Cell 144:16-26.

Lu, G., Henderson, D., Liu, L., Reinhart, P. H., Simon, S. A. 2005. TRPV1b: a functional human vanilloid receptor splice variant. Mol. Pharmacol. 67: 1119–1127.

McManus J., Graveley B. R. 2011. RNA structure and the mechanisms of alternative splicing. Current Opinion in Genetics & Development 21:373–379

Mata M., Munoz M. J., Mariano A., Pablo Fededa J., Schor I. E., Kornblihtt A. R. 2011. RNA Polymerase II Elongation at the Crossroads of Transcription and Alternative Splicing. Genetics Research 2011: 9.

Michaely, Tomchick P., Machius D. R., Anderson R.G. 2002. Crystal structure of a 12 ANK repeat stack from human ankyrinR. EMBO J. 21: 6387–6396.

Moiseenkova-Bell V. Y., Stanciu L. A., Serysheva I. I., Tobe B. J., Wensel T. G. 2008. Structure of *TRPV1* channel revealed by electron cryomicroscopy. Proc Natl Acad Sci USA 105: 7451–7455.

Moghe, P. V., Coger, R. N., Toner, M., Yarmush, M. L. 1997. Cell-cell interactions are essential for maintenance of hepatocyte function in collagen gel but not on matrigel. Biotechnol. Bioeng., 56: 706–711.

Modrek B., Lee,C. 2002. A genomic view of alternative splicing. Nature Genet. 30: 13–19.

Moroy, T., Heyd, F. 2007. The impact of alternative splicing *in vivo*: mouse models show the way. RNA 13: 1155–1171

Mudge J. M., Frankish A., Fernandez-Banet J. 2011. The origins, evolution, and functional potential of alternative splicing in vertebrates. Mol. Biol. Evolution. 28:2949–2959.

Nakae J., Kitamura T., Biggs III W. H., Arden K. C., Accili D. 2003. The forkhead transcription factor Foxo1 regulates adipocyte differentiation. Dev. Cell 4 :119–129.

Nern, A., Nguyen T.L., Herman T., Prakash S., Clandinin T. R., Zipursky L. S. 2005. An isoform-specific allele of Drosophila N-cadherin disrupts a late step of R7 targeting. Proc. Natl Acad. Sci. USA 102:12944–12949.

Neves G., Zucker J., Daly M., Chess A. 2004. Stochastic yet biased expression of multiple *Dscam* splice variants by individual cells. Nature Genet. 36: 240–246.

Nilsen, T. W. & Graveley, B. R. 2010. Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457–463.

Olshansky SJ, Carnes BA, Cassel C. In search of Methuselah: estimating the upper limits to human longevity. Science 1990;250: 634-40.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 40:1413–1415.

Parsons J. L., Dianova I. I., Khoronenkova S. V., Edelmann M. J., Kessler B. M., Dianov G. L. 2011. USP47 is a deubiquitylating enzyme that regulates base skip repair by controlling steadystate levels of DNA polymerase beta. Mol Cell 41: 609–615.

Patterson D., Cabelof D. C. 2011. Down syndrome as a model of DNA polymerase beta haploinsufficiency and accelerated aging, Mechanisms of Ageing and Development. 5.

Pingle S. C., Matta J. A., Ahern G. P. 2007. Capsaicin receptor: TRPV1 a promiscuous TRP channels. Handb Exp Pharacol 179: 155-71.

Rubin, C., Zody C. M., Eriksson J., Meadows R. S., Sherwood E., Webster M. L., Jiang L., Ingman M., Sharpe T., Ka S., Hallböök F., Besnier F., Carlborg O., Bed'hom B, Tixier-Boichard M., Jensen P., Siegel P., Lindblad-Toh K.,

Andersson L. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464: 587–591.

Sawaya M. R., Pelletier H., Kumar A., Wilson S. H., Kraut T. 1994. Crystal structure of rat DNA polymerase β: Evidence for a common polymerase mechanism. Science 264:1930–1935.

Schindler S., Szafranski K., Hiller M., Ali G. S., Palusa S. G., Backofen R., Platzer M., Reddy A. S. 2008. Alternative splicing at NAGNAG acceptors in *Arabidopsis thaliana* SR and SR-related protein-coding genes. BMC Genomics 9:159.

Schmucker D. Clemens C. J. Shu H., Worby A. C., Xiao J., Muda M., Dixon J. E., Zipursky L. S. 2000. Drosophila *Dscam* Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity. Cell 101: 671.

Schumacher M. A., Moff I., Sudanagunta S. P., Levine J. D. 2000. Molecular cloning of an N-terminal splice variant of the capsaicin receptor. Loss of N-terminal domain suggests functional divergence among capsaicin receptor subtypes. J. Biol. Chem. 272: 2756–2762.

Simonelli, V., D'Errico, M., Palli, D., Prasad, R., Wilson, S.H., Dogliotti, E. 2009. Characterization of DNA polymerase beta splicing variants in gastric cancer: The most frequent exon 2-deleted isoform is a non-coding RNA. Mutation Research 670: 79-87.

Skandalis A, Frampton M, Seger J, Richards M H: 2010. The adaptive significance of unproductive alternative splicing in primates. RNA 16:2014-2022.

Skandalis, A., Uribe, E., 2004. A survey of splice variants of the human hypoxanthine phosphoribosyl transferase and DNA polymerase beta genes: products of alternative or aberrant splicing? Nucleic Acids Research 32: 6557-6564.

Sorek R, Shamir R, Ast G. 2004. How prevalent is functional alternative splicing in the human genome? Trends Genet 20: 68–71.

Soergel D. A. W., Lareau L. F., Brenner S. E. 2006. Regulation of gene expression by the coupling of alternative splicing and nonsense mediated mRNA decay. Bioscience 175:196-211.

Sobol, R.W., Horton, J.K., Kuhn, R., Gu, H., Singhal, R.K., Prasad, R., Rajewsky, K., Wilson, S.H. 1996. Requirement of mammalian DNA polymerase-beta in base skip repair. Nature 379: 1853-186.

Srivastava D. K., Berg B. J., Prasad R., Molina J. T., Beard W. A., Tomkinson A. S., Wilson S. H. 1998. Mammalian abasic site base skip repair. Identification of reaction sequence and rate determining steps. J Biol Chem 273:21203-21209.

Srivastava D.K., Husain I., Arteaga C.L., Wilson S.H. 1999. DNA polymerase beta expression differences in selected human tumors and cell lines. Carcinogenesis 20: 1049–1054.

Starcevic D., Dalal S., Sweasy J. B. 2004. Is there a link between DNA polymerase beta and cancer? Cell Cycle 3: 998–1001.

Su A.I., Cooke M. P., Ching K. A., Hakak, Y., Walker J. R., Wiltshire T., Orth, A. P., Vega, R. G., Sapinoso L. M., Moqrich A. 2002. Large-scale analysis of the human and mouse transcriptomes. Proc.Natl. Acad. Sci. 99: 4465–4470.

Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. USA 101: 6062– 6067.

Sugnet C.W., Kent W.J., Ares M., and Haussler D. 2004. Transcriptome and genome conservation of alternative splicing events in humans and mice. Proc. 9th Pacific Symp. on Biocomputing 9: 66–77.

Suzuki M., Sato J., Kutsuwada K., Ooki G. Imai,M. 1999. Cloning of a stretch-inhibitable nonselective cation channel. J. Biol. Chem. 274: 6330-6335.

Sweasy J. B. 2003. Fidelity mechanisms of DNA polymerase beta. Prog Nucleic Acid Res Mol Biol 73:137–169.

Sweasy, J. B., Tieming L., Daniela S., Ka-Wai S., Char-Chang L., Daniel D., Shibani D. 2005. Expression of DNA polymerase beta mutants in mouse cells results in cellular transformation. Proc. Natl Acad. Sci. USA 102: 14350–14355.

Szallasi A., Blumberg P. M. 1999. Vanilloid (capsaicin) receptors and mechanisms. Pharmacol. Rev. 51:159–212

Tan X. H., Zhao M., Pan K. F., 2005. Frequent mutation related with overexpression of DNA polymerase beta in primary tumors and precancerous lesions of human stomach. Cancer Lett 220: 101-14.

Tazi, J., Bakkour, N., Stamm, S. 2009. Alternative splicing and disease. Biochim. Biophys. Acta 1792: 14–26.

Terenzi F, Ladd N. 2010. A Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. RNA Biol 7: 43-55.

Thompson T.E., Rogan,P.K., Risinger,J.I. and Taylor,J.A. 2002 Splice variants but not mutations of DNA polymerase beta are common in bladder cancer. Cancer Res 62: 3251–3256.

Tian W., Wang D. H., Cohen D. M. 2006. Regulation of TRPV1 by a novel renally expressed *TRPV1* splice variant. Am. J. Physiol. Renal Physiol. 290: F117–F126.

Tollervey J. R., Analysis of alternative splicing associated with aging and neurodegeneration in the human brain Genome Res. 2011 21: 1572-1582.

Tominaga M, Caterina M. J., Malmberg A. B., Rosen T. A., Gilbert H., Skinner K., Raumann B. E., Basbaum A. I., Julius D. 1998. Capsaicin receptor in the pain pathway. Neuron 21:531–543.

Tominaga M, Tominaga T. 2005. Structure and function of TRPV1. European Journal of Physiology. 451: 143–150.

Uchiyama Y, Takeuchi R, Kodera H, Sakaguchi K. 2009. Distribution and roles of X-family DNA polymerases in eukaryotes. Biochimie 91: 165–170.

Van Ommen. 2011. Induction of exon skipping in eukaryotic cells. Freepatents online.com.

Venables J. P., Vignal E., Baghdiguian S., Fort P., Tazi J. 2012. Tissue-Specific Alternative Splicing of Tak1 Is Conserved in Deuterostomes. Mol Biol Evol 29:

Wahl M. C, Will C. L, Luhrmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. Cell. 136:701–718.

Wagner E. J, Garcia-Blanco M. A. 2001. Polypyrimidine tract binding protein antagonizes exon definition. Mol Cell Biol 21: 3281-8.

Wang C, Hu H. Z., Colton C. K., Wood J. D., Zhu M. X. 2004. An alternative splicing product of the murine trpv1 gene dominant negatively modulates the activity of TRPV1 channels. *J Biol Chem* 279: 37423–30.

Wang E. T , Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore S. F, Schroth G. P, Burge B. C. 2008. Alternative isoform regulation in human tissue transcriptomes. Nature 456:470-476.

Xu Q, Modrek B, Lee C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res 30 :3754–66.

Yeo G, Holste D, Kreiman G, Burge CB. 2004. Variation in alternative splicing across human tissues. Genome Biol. 5:74.

Yu, Y. 2008. Dynamic regulation of alternative splicing by silencers that modulate 5’ splice site competition. Cell 135: 1224–123.

Zhou Z., Licklider L. J., Gygi S. P., Reed R. 2002. Comprehensive proteomic analysis of the human spliceosome. Nature 419: 182–185