Using Evolutionary Relationships of the *Poly-ADP-ribose Polymerase* (*PARP*) Gene Family to Identify Nearest Relative to the Undefined PARPs

By

Davia Denise Watkins

A thesis presented to the faculty at Middle Tennessee State University for partial fulfillment of the requirements of the Master of Science degree

Middle Tennessee State University
December 2022

Thesis Committee:

Dr. Rebecca Seipelt-Thiemann

Dr. David Nelson

Dr. James Robertson

Using Evolutionary Relationships of the *Poly-ADP-ribose Polymerase* (*PARP*) Gene Family to Identify Nearest Relative to the Undefined PARPs

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Rebecca Seipelt-Thiemann for being my mentor, guiding me through this process, helping me to strengthen my knowledge as a scientist, and providing wisdom and encouragement during these last few years, as being a graduate student was challenging. She ensured that I keep pushing and not give up, which gave me motivation to complete my master's degree and made me into a better graduate student. I would also like to thank Dr. David Nelson and Dr. James Robertson for being a part of my thesis committee and providing guidance for my project as their advice provided encouragement, challenged me to be more efficient with my writing skills, as well as to have a better understanding of general biological concepts.

I especially would like to thank my mother for all her love and support over these years and helping me to raise my son. Without my mother I would not be the woman I am today and would have not been afforded the opportunity to have an education, as she sacrificed time and time again so that I could earned my bachelor's and now master's degree. I would also like to thank my grandmother for her love and support, being there for me, and helping me raise my son.

ABSTRACT

Poly-ADP-ribose polymerases (PARPs) are proteins of a gene family that catalyze the post-translational modification (PTM) ADP-ribosylation. There are five human groups within the family that are defined according to known or unknown functions: DNA-PARPs, Tankyrase-PARPs, CCCH-PARPs, Macro-PARPs, and unknown-PARPs. Each of these members and their respective ortholog genes in four additional species (mouse, rat, zebrafish, and fruit fly) were studied to reveal evolutionary relationships and identify the nearest relative to each undefined member. This study showed PARP members group by function, that undefined-PARPs, 4, 6, 8, and 16 nearest relatives are DNA-PARPs, undefined-PARP10's nearest relative is CCCH or Macro-PARPs, and undefined-PARP11's nearest relative is CCCH-PARPs, and also provided evidence that another gene belongs in the *PARP* gene family, *ZC3HAV1L*. The evolution of Macro-PARP members 9, 14, and 15 revealed that gene duplication likely occurred.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In 2020, cancer was the deadliest disease with over 9 million deaths worldwide (World Health Organization (WHO) 2022). A person's genetic makeup, exposure to mutagenic chemicals or UV-radiation, viral infection, and spontaneous mutation contribute to cancer (WHO 2022). In 2020, there were a total of 2.26 million breast cancer cases and 685,000 breast cancer related deaths (WHO 2022). The genes *BRCA1* and *BRCA2* were identified as breast cancer-related genes in 1994 (*BRCA1*) and 1995 (*BRCA2*) and in 2020 were known to be involved in 3% of breast cancers and in 10% of ovarian cancers in women (Centers for Disease Control and Prevention (CDC) 2022 and Kwon et al. 2022). *BRCA1* and *BRCA2* encode proteins that normally suppress tumor formation because they maintain genomic integrity via assisting with double-stranded DNA-repair (Lord and Ashworth 2008).

However, when they are mutated they lose this function and permit genomic errors to accumulate, thus driving tumor formation and metastasis (Lord and Ashworth 2008). The promising area of research for BRCA-specific tumor therapy is the use of inhibitors of an enzyme called Poly-ADP-ribose polymerase (PARP), PARP inhibitors (Lord and Ashworth 2008). These inhibitors may specifically target the homologous recombination deficiency, which results from the *BRCA* mutations (Lord and Ashworth 2008). Indeed, mutation of either *PARP* or *BRCA* can be tolerated, but deletion of both is lethal (Lord and Ashworth 2008).

Poly-ADP-ribose Polymerases (PARPs) are proteins encoded by members of the *PARP* gene family that catalyze the post-translational modification ADP-ribosylation

(Daugherty et al. 2014, Li and Chen 2014, Vyas et al. 2013). PARP proteins use mono-ribose groups from nicotinamide adenine dinucleotide (NAD+) to generate either mono- or poly-ribose groups which are then added to amino acid residues (glutamic acid, lysine, and aspartic acid) of target or acceptor proteins (Li and Chen 2014, Vyas et al. 2013, Hassa and Hottiger 2008, Schreiber et al. 2006). Each PARP protein is defined by the presence of a 50 amino acid block that defines the PARP domain and each exhibits mono- (MAR), poly- (PAR), or no ribosylase activity (Li and Chen 2014, Hassa and Hottiger 2008, Schreiber et al. 2006, Vyas et al. 2013, Nguewa et al. 2005). The addition of MAR or PAR to target proteins modulates a variety of cellular processes including DNA repair, chromatin remodeling, transcription, apoptosis, mitosis, ubiquitination, proteasomal degradation, protein-protein interactions, viral RNA replication, cellular proliferation, and cancer (Li and Chen 2014, Vyas et al. 2013, Daugherty et al. 2014).

Gene families such as the *PARP* gene family are a group of genes originating from a single ancestral gene. In gene family development, an ancestral gene along with its regulatory sequences is duplicated in either a whole genome or gene duplication and then each copy is subject to mutation and divergence (Andersson et al. 2015, Fig. 1). Outcomes of divergence among members of a gene family include conservation of gene function (a gene retains the original function), pseudogenization (one gene copy is no longer transcribed or becomes a non-functional pseudogene due to a mutation), sub-functionalization (an original gene copy had multiple functions that were split between duplicated copies) and neo-functionalization (one gene copy obtains a new function via mutation) (Magadum et al. 2013, Lynch and Conery 2000, Zhang 2003, Ohno 1970,

Hurles 2004, Fig. 1). Over time, repeated duplication and divergence generate larger gene families whose members, sequences, functions, and gene expression have diverged from the ancestral gene (Andersson et al. 2015). Duplications that occurred more recently will have more conservation than duplications that occurred long ago (Maynard et al. 2014).

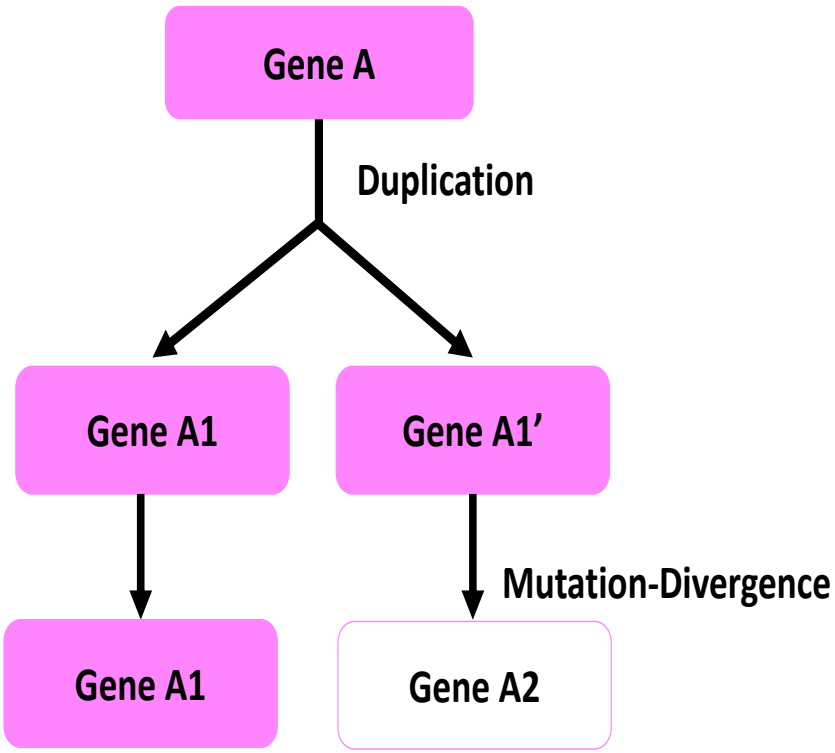


Figure. 1. Gene Duplication and Divergence. A single gene, which is noted as Gene A, can be duplicated in a gene or genome duplication event, producing two copies (Gene A1 and Gene A1'). Mutations accumulate over time to one or both copies creating a divergence between the two genes (Genes A1, A2) in structure, sequence, location, expression, and/or function. This figure is adapted from Andersson et al. (2015).

In order to learn more about the origin and development of this gene family, the first goal of this study was to use multiple lines of evidence to provide a comprehensive view of the evolutionary relatedness among the members of the *PARP* gene family. PARP homologs are found in over 200 species, many with multiple gene family members (Cunningham et al. 2022). A previous study conducted by Citarelli et al. (2010) focused on the evolutionary history of eukaryotic *PARP* gene family members, along with PARP ancestral proteins. This study revealed that PARP proteins belonged to six clade groupings, including that of human known PARPs and undefined-PARPs, and also identified two suggested ancestral genes. A study conducted by Perina et al. (2014) also discusses these six groupings in similar detail across human and other species.

In our study, we focus on identifying which human undefined-PARPs belong to known PARP groups, by incorporating other model species that have orthologous *PARP* genes, such as mouse, rat, zebrafish, and fruit fly to solidify findings, and associate our findings with previously identified groupings to analyze the evolvement of members within this gene family. The five species used in this study (Table 1) were chosen because they have well-annotated genomes. The PARP family has been most well-studied in human, where there are seventeen members that have been categorized into five sub-groups based on function and the presence of specific protein domains in addition to their PARP domain (Table 2). Since categorization of genes by function is one line of evidence regarding the relatedness of these genes, we have started with a brief overview of the human gene family in this study.

Table. 1. PARP Homologs in Human, Mouse, Rat, Zebrafish, and Fruit Fly (NCBI 2022, Kent et al. 2002, Bult et al. 2019, Smith et al. 2019, Cunningham et al. 2022, and Yates et al. 2022)

Human	Mouse	Rat	Zebrafish	Fruit Fly
PARP1	*PARP1*	*PARP1*	*PARP1*	*PARP*
PARP2	*PARP2*	*PARP2*	*PARP2*	NA
PARP3	*PARP3*	*PARP3*	*PARP3*	NA
PARP4	*PARP4*	*PARP4*	*PARP4*	NA
PARP5a	*PARP5a*	*PARP5a*	*TNKSa*(*PARP5_1*) *TNKSb*(*PARP5_2*)	*TNKS*
PARP5b	*PARP5b*	*PARP5b*	NA	*TNKS*
PARP6	*PARP6*	*PARP6*	*PARP6a* *PARP6b*	NA
PARP7	*PARP7*	*PARP7*	*PARP7*	NA
PARP8	*PARP8*	*PARP8*	*PARP8*	NA
PARP9	*PARP9*	*PARP9*	*PARP9*	NA
PARP10	*PARP10*	*PARP10*	*PARP10*	NA
PARP11	*PARP11*	*PARP11*	*PARP11*	NA
PARP12	*PARP12*	*PARP12*	*PARP12a* *PARP12b*	NA
PARP13	*PARP13*	*PARP13*	NA	NA
PARP14	*PARP14*	*PARP14*	*PARP14_1* (*PARP14rs1*) *PARP14rs2.1* *PARP14_2* (*PARP14rs3*) *PARP14_3* (*PARP14rs4*)	
PARP15	NA	NA	NA	NA
PARP16	*PARP16*	*PARP16*	*PARP16*	*PARP16*

Table. 2. Human *PARP* Gene and Protein Information. Gene name, sub-group, chromosome locations, exon count, protein size, protein domains, and ADP-ribosylation activity.

Group[1,2]	Gene	Chromosome[3]	Exon Count[4]	Protein size (aa)[4]	Protein Domains[5-12]	ADP-ribosylation[5]
DNA-dependent-PARPs	*PARP1*	1q42.12	23	1014	ZINC FINGER, BRCT, WGR, PARP	Poly
	PARP2	14q11.2	16	570	WGR, PARP	Poly
	PARP3	3p21.2	11	533	WGR, PARP	Poly
Tankyrase-PARPs	*PARP5a*	8p23.1	27	1327	ANK, SAM, PARP	Poly
	PARP5b	10q23.32	27	1166	ANK, SAM, PARP	Poly
CCCH-PARPs	*PARP7*	3q25.31	6	657	CCCH, WWE, PARP	Mono
	PARP12	7q34	12	701	CCCH, WWE, PARP	Mono
	PARP13	7q34	13	902	CCCH, WWE, PARP	None
Macro-PARPs	*PARP9*	3q21.1	11	819	MACRO, PARP	None
	PARP14	3q21.1	17	1801	MACRO, WWE, PARP	Mono
	PARP15	3q21.1	12	678	MACRO, PARP	Mono
Other-PARPs	*PARP4*	13q12.12	34	1724	BRCT, PARP	Poly
	PARP6	15q23	24	630	PARP	Mono
	PARP8	5q11.1	26	854	PARP	Mono
	PARP10	8q24.3	11	1025	PARP	Mono
	PARP11	12p13.32	8	338	WWE, PARP	Mono
	PARP16	15q22.31	6	322	PARP	Mono

[1] Li and Chen 2014
[2] Vyas et al. 2013
[3] NCBI 2021
[4] UCSC Genome Browser 2002
[5] Li and Chen 2014
[6] Sterile Alpha Motif
[7] Cys-Cys-Cys-His
[8] BRCA1 C Terminus
[9] Ankyrin repeat
[10] Try-Gly-Arg
[11] Try-Try-Glu
[12] Poly-ADP-ribose polymerase

The first group is the DNA-PARPs, which includes PARPs 1, 2, and 3, were first identified as DNA binding proteins (Vyas et al. 2013). The PARP domain for each of these family members exhibits PAR activity (Li and Chen 2014). Each of these PARPs also contains a WGR domain (tryptophan, glycine, and arginine amino acids) (Li et al. 2014) which has been suggested to be a nucleic acid-binding domain (Letunic et al. 2021). When DNA strand breaks occur due to DNA damage, each of these PARPs can recognize the breaks and assist with DNA repair (Li and Chen 2014, Beck et al. 2014, Gibson and Kraus 2012, Hakmé et al. 2008). PARP1, has additional domains such as the Zinc finger domain and the BRCA1 C terminus (BRCT) domain (Li and Chen 2014). The Zinc finger domain can bind DNA, while the BRCT domain is a phospho-peptide binding domain involved with DNA-damage control and is found in proteins involved with DNA repair and cell cycle activity (Letunic et al. 2021, Li and Chen 2014, Glover et al. 2004, Yu et al. 2003). PARP1 is vital for base excision repair (BER) and single-stranded break repair (SSBR), as it recruits and binds to the XRCC1 protein to assist with damaged areas of DNA (Li and Chen 2014, Parsons et al. 2005). PARP1 also facilitates double-stranded break repairs (DSBR), as DSBR proteins Ku and DNA-PKcs interaction with PARP1 is necessary for non-homologous end joining repair activity (Kraus 2008, Li et al. 2004, Li and Chen 2014, Ruscetti et al. 1998). PARP2 is also important for BER and SSBR activity, while PARP3 has been found to have a functional role in DSBR activity (Beck et al. 2014 and Li and Chen 2014).

The second group of PARPs is the Tankyrase-PARPs (PARPs 5a, 5b) which exhibit PAR activity and also contain ankyrin repeat (ANK), sterile alpha motif (SAM), and PARP domains (Li and Chen 2014). The ANK domain contains ankyrin repeated regions that appear in helix loop formations and are involved in protein-protein interactions (Letunic et al. 2021). The SAM domain is associated with protein binding, cell signaling, and RNA-binding activity (Letunic et al. 2021). These PARPs regulate protein degradation, such as the proteasome-based degradation of the telomeric repeat binding factor 1 (TRF1) protein which modulates extension of telomeres via PAR activity (Smith et al. 1998, Smith and de Lange 2000, Li and Chen 2014, NCBI 2022). PARP5a can cause protein degradation of the 3BP2 protein via PAR activity, which can allow for normal bone development (Levaot et al. 2011, Li and Chen 2014). Tankyrase-PARPs have also been proposed targets for cancer therapy treatments as they modulate proteosome-based degradation of the axin protein (PAR activity), which turns on the Wnt signaling pathway (Riffell et al. 2012, Huang et al. 2009, Li and Chen 2014). Tankyrase-PARP inhibitor (XAV939) was able to hinder the Wnt signaling pathway and allow stabilization of both axin1 and 2 proteins (Li and Chen 2014, Huang et al. 2009). PARP5a also regulated the nuclear mitotic apparatus protein (NUMA) via PAR activity to control cohesion of sister chromatids (Li and Chen 2014, Chang et al. 2005).

The third group of PARPs is known as the zinc finger CCCH (Cys-Cys-Cys-His)-PARPs (PARP7, 12, and 13) and contain the CCCH, WWE, and PARP domains (Li and Chen 2014, NCBI 2021, He et al. 2012, Vyas et al. 2013). PARP7 and PARP12 have (MAR) activity, while PARP13 does not have ribosylase activity (Li and Chen 2014). The WWE

domain, which is comprised of tryptophan and glutamic acid amino acids (He et al. 2012) is involved with proteasome degradation and ubiquitination (Li and Chen 2014, Letunic et al. 2021). CCCH-zinc finger proteins, containing the CCCH-domain, bind RNA and play a role in RNA metabolism (Hajikhezri et al. 2020). These PARPs modulate transcription and viral (RNA) replication (Li and Chen 2014, Daugherty et al. 2014, Vyas et al. 2013). For example, PARP7 repressed transcription of the aryl hydrocarbon receptor (AHR) (MacPherson et al. 2013, Li and Chen 2014) and PARP12 regulated replication of the Venezuelan equine encephalitis virus (VEEV) (Atasheva et al. 2012, Li and Chen 2014). A PARP13 isoform interacted with viral RNA and inhibited its accumulation (Li and Chen 2014, Gao et al. 2002). Finally, overexpression of PARP13 limited viral replication of virus families, including retrovirus (murine leukemia virus), filoviruses (Ebola and Marburg), togavirus (Sindbis), as well the hepadnavirus (Hepatitis B virus) (Daugherty et al. 2014, Gao et al. 2002, Muller et al. 2007, Bick et al. 2003, Mao et al. 2013).

The fourth PARP group is the Macro-PARPs (PARP9, 14, and 15) which contain both a Macro and PARP domain (Li and Chen 2014, NCBI 2021, Vyas et al. 2013). Macro domain proteins regulate transcription, DNA repair, remodeling of chromatin, and protein interactions (Letunic et al. 2021, Han et al. 2011). Proteins with macro domains have been localized to DNA single and double stranded breaks and areas that underwent DNA repair (Han et al. 2011, Timinszky et al. 2009, Gottschalk et al. 2009, Ahel et al. 2009). Both PARP14 and PARP15 exhibit MAR activity, while PARP9 does not exhibit ribosylase activity (Li and Chen 2014). PARP14 alone has a WWE domain similar

to the CCCH-PARPs (Li and Chen 2014). Both PARP9 and 14 have been found to regulate transcription, as PARP9 modulated transcription in diffuse B-cell lymphoma cells and PARP14 modulated transcription of signal transducer and activator of transcription 6 (STAT6) (Li and Chen 2014, Juszczynyski et al. 2006, Mehrotra et al. 2011, NCBI 2022). PARP9 functions in the DNA damage response by regulating ubiquitination of DSBR proteins 53BP1 and BRCA1 (Li and Chen 2014, Yan et al. 2013). The function of PARP15 is not yet known (Li and Chen 2014) but exhibited transcriptional repression in a genetic screen for transcription regulators in human embryonic kidney 293 (HEK293) cells (Han et al. 2011, Aguiar et al. 2005).

The fifth group is the "other" or undefined-PARPs (PARP4, 6, 8, 10, 11, and 16) (Li and Chen 2014, NCBI 2021, Vyas et al. 2013). All undefined-PARPs have (MAR) activity except PARP4, which has (PAR) activity (Li and Chen 2014). They share only the PARP domain and have various functions including response to DNA damage, regulatory cell cycle activity, anti-viral activity, and protein degradation (Li and Chen 2014, Guo et al. 2019). PARP4 contains a BRCT domain, as similarly seen in PARP1 (Li and Chen 2014). Mice deficient in PARP4 activity showed an increase in tumors caused by mutagens, indicating that PARP4 may play a role in the suppression of tumor formation (Li and Chen 2014, Raval-Fernandes et al. 2005). PARP6 modulated cellular proliferation, while PARP10 interacted with known proto-oncogene *Myc* and inhibited the oncogenic *c-Myc* (Li and Chen 2014, Tuncel et al. 2012, Yu et al. 2005). PARP16 regulated protein levels of karyopherin-Beta1, a nuclear import protein (Li and Chen 2014, Di Paola et al. 2012, Carden et al. 2018). Additionally, when endoplasmic reticulum cellular stress occurred,

PARP16 ribosylated and activated proteins PERK and IRE1alpha (Jwa and Chang 2012, Li and Chen 2014). PARP8 performs (MAR) activity (Li and Chen 2014). PARP11 contains a WWE domain and uses its ribosylase activity to mediate viral responses in herpes simplex-1, vesicular stomatitis, and influenza A virus (Li and Chen 2014, Guo et al. 2019). Additionally, PARP11 was found to modulate the anti-viral activity of type I interferon (IFN-I) protein via MAR activity of Beta-TrCP (ubiquitin E3 ligase β-transducin repeat) protein, and MAR activity of this protein allowed for IFN alpha and beta receptor subunit 1 (IFNAR1) to be ubiquitinated and degraded (Guo et al. 2019). The placement of these undefined-PARPs with a sub-group is the second part of this project.

Conservation of function and domains among the gene family members and among the distinct groups necessarily depends on the conservation of particular amino acids. However, other data can be used to further refine these broad relationships. In this study we investigated additional lines of valid evolutionary-based evidence including amino acid identity and similarity across the entirety of the proteins, conservation of gene expression, and conservation of genomic context (Sievers et al. 2011, Stothard 2000, NCBI 2022, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Kent et al. 2002, Obayashi et al. 2019, Loraine et al. 2015, GTEx 2021-2022, Tostivint et al. 2014, Ocampo-Daza et al. 2011, Kolishovski et al. 2019, Maynard et al. 2014). Using these combined data, we provide a comprehensive view of the relationships among the gene family and identify the likely nearest neighbor for the "undefined" group of PARPs.

METHODS

Species of Study

The species used in this study were chosen because they have genomes that have been well-studied and are better annotated for gene structure, sequence, and gene function than many other species. These species are human (*Homo sapiens*), mouse (*Mus musculus*), rat (*Rattus norvegicus*), zebrafish (*Danio rerio*), and fruit fly (*Drosophila melanogaster)* (The Alliance 2019, NCBI 2022, Kent et al. 2002, Cunningham et al. 2022, Bult et al. 2019, Smith et al. 2019, Yates et al. 2022).

The human *PARP* gene family has seventeen annotated members, the mouse and rat family contain sixteen members, the zebrafish family contains twenty members, and the fruit fly family has three members all totaling to 72 species (NCBI 2022, Kent et al. 2002, Cunningham et al. 2022, Bult et al. 2019, Smith et al. 2019, Yates et al. 2022; Table 1, Table 2).

PARP homolog identification and sequence sources

The HUGO Gene Nomenclature Committee (HGNC) was used to identify the seventeen human *PARP* gene family members and sixteen mouse orthologs (Braschi et al. 2019). The University of California Santa Cruz (UCSC) Genomic Browser was used to download human exon and protein sequences (Kent et al. 2002; Table 3). Human gene (*Homo sapiens*) reference sequences were identified using the MANE (NCBI and Ensembl combined annotated sequences) Transcript Project (v 0.92) except *PARP9*, which does not have a MANE transcript, so the curated NCBI reference sequence was utilized (Kent et al. 2002). The mouse genome informatics (MGI) database was utilized to retrieve mouse (*Mus musculus*) PARP protein sequences (Bult et al. 2019; Table 4). Rat (*Rattus norvegicus*) protein sequences were retrieved from the Rat Genome Database (RGD) and NCBI (Smith et al. 2019, NCBI 2022; Table 5). Fruit fly (*Drosophila melanogaster*) and zebrafish (*Danio rerio*) protein sequences were obtained from the Ensembl Main and Ensembl Metazoan databases (Cunningham et al. 2022 and Yates et al. 2022; Table 6). *ZC3HAV1L* in human, mouse, and rat protein sequences were obtained from Ensembl Main, MGI, and RGD databases (Cunningham et al. 2022, Bult et al. 2019, Smith et al. 2019; Tables 3-5).

Table. 3. Accession numbers of the Human members of the *PARP* Family. The human sequences were obtained from the UCSC and Ensembl Main databases (Kent et al. 2002, Cunningham et al. 2022).

Species	Gene Name	Protein ID# and names	Gene ID #
Human	*PARP1*	ENST00000366794.10	ENST00000366794.10
Human	*PARP2*	ENST00000429687.8	ENST00000429687.8
Human	*PARP3*	ENST00000398755.8	ENST00000398755.8
Human	*PARP4*	ENST00000381989.4	ENST00000381989.4
Human	*PARP5a*	ENST00000310430.11	ENST00000310430.11
Human	*PARP5b*	ENST00000371627.5	ENST00000371627.5
Human	*PARP6*	ENST00000569795.6	ENST00000569795.6
Human	*PARP7*	ENST00000295924.12	ENST00000295924.12
Human	*PARP8*	ENST00000281631.10	ENST00000281631.10
Human	*PARP9*	NP_001139576.1	NM_001146102.2
Human	*PARP10*	ENST00000313028.12	ENST00000313028.12
Human	*PARP11*	ENST00000228820.9	ENST00000228820.9
Human	*PARP12*	ENST00000263549.8	ENST00000263549.8
Human	*PARP13*	ENST00000242351.10	ENST00000242351.10
Human	*ZC3HAV1L*	ZC3HAV1L-201 peptide.prot	NA
Human	*PARP14*	ENST00000474629.7	ENST00000474629.7
Human	*PARP15*	ENST00000464300.7	ENST00000464300.7
Human	*PARP16*	ENST00000649807.2	ENST00000649807.2

Table. 4. Accession numbers of the Mouse members of the *PARP* Family. The mouse sequences were obtained from the MGI database (Bult et al. 2019).

Mouse	*PARP1*	P11103	NA
Mouse	*PARP2*	O88554	NA
Mouse	*PARP3*	Q3ULW8	NA
Mouse	*PARP4*	ENSMUSP00000124258.1	NA
Mouse	*PARP5a*	Q6PFX9	NA
Mouse	*PARP5b*	ENSMUSP00000025729.5	NA
Mouse	*PARP6*	ENSMUSP00000026267.8	NA
Mouse	*PARP7*	Q8C1B2	NA
Mouse	*PARP8*	ENSMUSP00000022239.6	NA
Mouse	*PARP9*	ENSMUSP00000110528.1	NA
Mouse	*PARP10*	Q8CIE4	NA
Mouse	*PARP11*	ENSMUSP00000036127.5	NA
Mouse	*PARP12*	Q8BZ20	NA
Mouse	*PARP13*	ENSMUSP00000110550.1	NA
Mouse	*ZC3HAV1L*	ENSMUSP00000062475.3	NA
Mouse	*PARP14*	ENSMUSP00000037657.8	NA
Mouse	*PARP16*	Q7TMM8	NA

Table. 5. Accession numbers of the Rat members of the *PARP* Family. The rat sequences were obtained from the RGD and NCBI databases (Smith et al. 2019, NCBI 2022).

Rat	*PARP1*	NP_037195	NA
Rat	*PARP2*	NP_001099500	NA
Rat	*PARP3*	NP_001008329	NA
Rat	*PARP4*	XP_038949775	NA
Rat	*PARP5a*	NP_001099554	NA
Rat	*PARP5b*	NP_001101077	NA
Rat	*PARP6*	NP_001100298	NA
Rat	*PARP7*	NP_001101149	NA
Rat	*PARP8*	XP_038959446	NA
Rat	*PARP9*	NP_001096821	NA
Rat	*PARP10*	XP_006241972	NA
Rat	*PARP11*	XP_038964758	NA
Rat	*PARP12*	XP_006236408	NA
Rat	*PARP13*	NP_766633	NA
Rat	*ZC3HAV1L*	XP_342662	NA
Rat	*PARP14*	NP_001178588	NA
Rat	*PARP16*	NP_001014115	NA

Table. 6. Accession numbers of the Zebrafish and Fruit fly members of the *PARP* Family. The zebrafish and fruit fly sequences were obtained from the Ensembl Main and Ensembl Metazoan databases (Cunningham et al. 2022, Yates et al. 2022).

Zebrafish	*PARP1*	PARP1-201peptide.prot	NA
Zebrafish	*PARP2*	PARP2-201peptide.prot	NA
Zebrafish	*PARP3*	PARP3-201peptide.prot	NA
Zebrafish	*PARP4*	PARP4-201peptide.prot	NA
Zebrafish	*TNKSa*	TNKSa-201peptide.prot	NA
Zebrafish	*TNKSb*	TNKSb-201peptide.prot	NA
Zebrafish	*PARP6a*	PARP6a-201peptide.prot	NA
Zebrafish	*PARP6b*	PARP6b-201peptide.prot	NA
Zebrafish	*PARP7*	TIPARP-201peptide.prot	NA
Zebrafish	*PARP8*	PARP8-201peptide.prot	NA
Zebrafish	*PARP9*	PARP9-201peptide.prot	NA
Zebrafish	*PARP10*	si:ch1073-296d18.1(si9.prot)	NA
Zebrafish	*PARP11*	PARP11-201peptide.prot	NA
Zebrafish	*PARP12a*	PARP12a-201peptide.prot	NA
Zebrafish	*PARP12b*	PARP12b-201peptide.prot	NA
Zebrafish	*PARP14rs1*	PARP14_1Dre si:ch211-219a4.3 PARP14rs1 (si20.prot)	NA
Zebrafish	*PARP14rs3*	PARP14_2Dre CABZ01079480.1 PARP14rs3 (CABZ01079480.1-201 peptide.prot)	NA
Zebrafish	*PARP14rs4*	PARP14_3Dre PARP14rs4 (FO744833.2-201 peptide.prot)	NA
Zebrafish	*PARP16*	PARP16-201peptide.prot	NA
Fruit Fly	*PARP1*	PARP-RB peptide.prot	NA
Fruit Fly	*TNKS*	TNKS-RA peptide.prot	NA
Fruit Fly	*PARP16*	PARP16-RA peptide.prot	NA

Sequence conservation and Phylogenetic analyses using amino acids

The Clustal Omega multiple sequence alignment tool was used to align the PARP and protein sequences by similar sequence patterns between family members in each species and across species (Sievers et al. 2011). This then allowed the generation of an output FASTA sequence protein file from Clustal Omega of these members to use for pairwise identity and similarity sequence manipulation analysis (Stothard 2000). Pairwise similarity and identity amino acid conservation percentages of PARP protein family members across species was obtained with the use of the sequence manipulation suite (SMS) tool (Stothard 2000). These percentages were then downloaded and re-organized using a custom PERL Script. Text files were colored in excel. The Clustal Omega and MView multiple sequence alignment tools were used to align the PARP protein sequences by similar sequence patterns between family members (Sievers et al. 2011, Brown et al. 1998). MView color coded these sequences in a box color shaded form to show which amino acids are identical or divergent from one another (Brown et al. 1998, Brown 1999). The COBALT multiple sequence alignment tool was similarly used to align protein sequences relative to sequence patterns between members (Papadopoulos and Agarwala 2007, NCBI 2022).

Molecular Evolutionary Genetics Analysis (MEGA) was utilized for the phylogenetic tree analysis of the PARP protein sequences of each individual species and all species as one analysis (Kumar et al. 2018, Tamura et al. 2021, Stecher et al. 2020). The Neighbor-Joining tree statistical method was utilized for phylogeny bootstrapping purposes so that the members that were most closely related were

displayed in the tree results according to a replicated number of 500 or out of 500 replicates (Felsenstein 1985, Saitou and Nei 1987). The Poisson Correction method was used to quantify the evolutionary distances between members (Zuckerkandl and Pauling 1965) and the Pairwise Deletion option was chosen to remove amino acid positions within each pair of sequences that do not show major inference on the evolution that has taken place between the sequences (Zuckerkandl and Pauling 1965). There were three tree threads used for each analysis. The Maximum-Likelihood tree statistical method along with the JTT-matrix based model was utilized for phylogeny bootstrapping purposes so that the members that were most closely related amongst all species were displayed in tree results according to a log mathematical measurement of likelihood and produced a total number of 100 replicates (Felsenstein 1985, Jones et al. 1992). There were also three tree threads used for analysis.

Gene expression analyses

Pairwise co-expression was used to determine the similarity of gene expression level of each pair of the gene family for human, mouse, rat, zebrafish, and fruit fly members using the COXPRESdb database (Obayashi et al. 2019). Gene expression similarity was scaled by color using the Entrez gene ID numbers with the Hcluster software on the COXPRESdb database (Obayashi et al. 2019). Normalized human transcriptome gene expression data (transcript per million-TPM) for 54 normal human tissues were obtained from the Genotype-Tissue Expression (GTEx) database (GTEx 2021) and were used to investigate overall gene family clustering of gene expression levels using R Studio Windows version 4.0.5. R Studio packages and R plots used to

analyze differential gene expression included the edgeR Package to generate multi-dimensional scale (MDS) plots, and cluster dendrogram plots (Loraine et al. 2015 and Robinson et al. 2010). The GTEx database was also used to generate a gene-specific heat map of *PARP* genes expression by tissue (GTEx 2022).

Gene order and Synteny analyses

Majority of the *PARP* genes for all five species were located within its genome using the NCBI and Ensembl databases (NCBI 2022, Cunningham et al. 2022, Yates et al. 2022). Neighboring genes and direction of transcription was identified, and each were compared to identify gene duplications, inversions, and gene loss (NCBI and Ensembl, Tables 7-11). Synteny analyses involve the study of genomic ancestry of species to determine if gene order across species chromosomes has been conserved over time (JAX Synteny Browser, Kolishovski et al. 2019). Synteny for human-mouse in the regions of each *PARP* gene was performed using the JAX Synteny Browser (Kolishovski et al. 2019).

.

Table. 7. Human *PARP* Chromosomal Context NCBI Internet URL Links. Each of these Human *PARP* gene members in each species chromosome context can be searched on NCBI. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP1 Human	https://www.ncbi.nlm.nih.gov/gene/142
PARP2 Human	https://www.ncbi.nlm.nih.gov/gene/10038
PARP3 Human	https://www.ncbi.nlm.nih.gov/gene/10039
PARP4 Human	https://www.ncbi.nlm.nih.gov/gene/143
PARP5a Human	https://www.ncbi.nlm.nih.gov/gene/8658
PARP5b Human	https://www.ncbi.nlm.nih.gov/gene/80351
PARP6 Human	https://www.ncbi.nlm.nih.gov/gene/56965
PARP7 Human	https://www.ncbi.nlm.nih.gov/gene/25976
PARP8 Human	https://www.ncbi.nlm.nih.gov/gene/79668
PARP9 Human	https://www.ncbi.nlm.nih.gov/gene/83666
PARP10 Human	https://www.ncbi.nlm.nih.gov/gene/84875
PARP11 Human	https://www.ncbi.nlm.nih.gov/gene/57097
PARP12 Human	https://www.ncbi.nlm.nih.gov/gene/64761
PARP13 Human	https://www.ncbi.nlm.nih.gov/gene/56829
PARP14 Human	https://www.ncbi.nlm.nih.gov/gene/54625
PARP15 Human	https://www.ncbi.nlm.nih.gov/gene/165631
PARP16 Human	https://www.ncbi.nlm.nih.gov/gene/54956

Table. 8. Mouse *PARP* Chromosomal Context NCBI Internet URL Links. Each of these Mouse *PARP* gene members in each species chromosome context can be searched on NCBI. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP1 Mouse	https://www.ncbi.nlm.nih.gov/gene/11545
PARP2 Mouse	https://www.ncbi.nlm.nih.gov/gene/11546
PARP3 Mouse	https://www.ncbi.nlm.nih.gov/gene/235587
PARP4 Mouse	https://www.ncbi.nlm.nih.gov/gene/328417
PARP5a Mouse	https://www.ncbi.nlm.nih.gov/gene/21951
PARP5b Mouse	https://www.ncbi.nlm.nih.gov/gene/74493
PARP6 Mouse	https://www.ncbi.nlm.nih.gov/gene/67287
PARP7 Mouse	https://www.ncbi.nlm.nih.gov/gene/99929
PARP8 Mouse	https://www.ncbi.nlm.nih.gov/gene/52552
PARP9 Mouse	https://www.ncbi.nlm.nih.gov/gene/80285
PARP10 Mouse	https://www.ncbi.nlm.nih.gov/gene/671535
PARP11 Mouse	https://www.ncbi.nlm.nih.gov/gene/101187
PARP12 Mouse	https://www.ncbi.nlm.nih.gov/gene/243771
PARP13 Mouse	https://www.ncbi.nlm.nih.gov/gene/78781
PARP14 Mouse	https://www.ncbi.nlm.nih.gov/gene/547253
PARP15 Mouse	NA
PARP16 Mouse	https://www.ncbi.nlm.nih.gov/gene/214424

Table. 9. Rat *PARP* Chromosomal Context NCBI Internet URL Links. Each of these Rat *PARP* gene members in each species chromosome context can be searched on NCBI. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP1 Rat	https://www.ncbi.nlm.nih.gov/gene/25591
PARP2 Rat	https://www.ncbi.nlm.nih.gov/gene/290027
PARP3 Rat	https://www.ncbi.nlm.nih.gov/gene/300985
PARP4 Rat	https://www.ncbi.nlm.nih.gov/gene/361046
PARP5a Rat	https://www.ncbi.nlm.nih.gov/gene/290794
PARP5b Rat	https://www.ncbi.nlm.nih.gov/gene/309512
PARP6 Rat	https://www.ncbi.nlm.nih.gov/gene/300759
PARP7 Rat	https://www.ncbi.nlm.nih.gov/gene/310467
PARP8 Rat	https://www.ncbi.nlm.nih.gov/gene/294762
PARP9 Rat	https://www.ncbi.nlm.nih.gov/gene/303905
PARP10 Rat	https://www.ncbi.nlm.nih.gov/gene/102548778
PARP11 Rat	https://www.ncbi.nlm.nih.gov/gene/500323
PARP12 Rat	https://www.ncbi.nlm.nih.gov/gene/362343
PARP13 Rat	https://www.ncbi.nlm.nih.gov/gene/252832
PARP14 Rat	https://www.ncbi.nlm.nih.gov/gene/303903
PARP15 Rat	NA
PARP16 Rat	https://www.ncbi.nlm.nih.gov/gene/315760

Table. 10. Zebrafish *PARP* Chromosomal Context NCBI-Ensembl Internet URL Links. Each of these Zebrafish *PARP* gene members in each species chromosome context can be searched on NCBI and Ensembl. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP1 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/560788 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000019529;r=20:43648369-43689779
PARP2 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/100007906 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000079202;r=2:37836821-37846849
PARP3 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/335495 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000003961;r=6:42038358-42049350;t=ENSDART00000022949
PARP4 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/558045 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000069934;r=9:20483846-20502622
PARP5 Zebrafish Tnksa	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000077650;r=21:19925910-20108849 https://www.ncbi.nlm.nih.gov/gene/559022
PARP5 Zebrafish Tnksb	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000062897;r=5:8964926-9027936 https://www.ncbi.nlm.nih.gov/gene/567533
PARP6a Zebrafish	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?db=core;g=ENSDARG00000042703;r=18:887122-898873 https://www.ncbi.nlm.nih.gov/gene/436810
PARP6b Zebrafish	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?db=core;g=ENSDARG00000070473;r=25:29099477-29134654 https://www.ncbi.nlm.nih.gov/gene/100002364

Table. 10. Zebrafish *PARP* Chromosomal Context Continued. Each of these Zebrafish *PARP* gene members in each species chromosome context can be searched on NCBI and Ensembl. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP7 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/563506 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000061841;r=18:34599315-34642086
PARP8 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/568910 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000059789;r=5:40844968-40910880
PARP9 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/799852 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000006848;r=9:23754275-23765480
PARP10 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/564886 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000087145;r=19:11952309-11977783
PARP11 Zebrafish	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000098716;r=4:71989418-72002734;t=ENSDART00000170996
PARP12a Zebrafish	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?db=core;g=ENSDARG00000042496;r=18:12839126-12858016;t=ENSDART00000130343 https://www.ncbi.nlm.nih.gov/gene/567195
PARP12b Zebrafish	https://uswest.ensembl.org/Danio_rerio/Gene/Summary?db=core;g=ENSDARG00000057974;r=4:9813753-9822519 https://www.ncbi.nlm.nih.gov/gene/553633
PARP13 Zebrafish	NA

Table. 10. Zebrafish *PARP* Chromosomal Context Cont. Each of these Zebrafish *PARP* gene members in each species chromosome context can be searched on NCBI and Ensembl. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP14rs1 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/791754 https://zfin.org/ZDB-TSCRIPT-090929-9946 Note: Protein Sequence was originally obtained from Ensembl
PARP14rs3 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/100148704 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000098398;r=9:56443416-56459826
PARP14rs4	https://www.ncbi.nlm.nih.gov/gene/562648 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000112454;r=10:21913105-21955231
PARP15 Zebrafish	NA
PARP16 Zebrafish	https://www.ncbi.nlm.nih.gov/gene/405842 https://uswest.ensembl.org/Danio_rerio/Gene/Summary?g=ENSDARG00000018593;r=25:31868268-31882618;t=ENSDART00000022325

Table. 11. Fruit Fly *PARP* Chromosomal Context NCBI-Ensembl Internet URL Links. Each of these Fruit Fly *PARP* gene members in each species chromosome context can be searched on NCBI, Ensembl Main, and the Ensembl Metazoan database. The following URL links were used to begin context analysis. The gene id numbers are listed at the end of the link.

PARP1 Fruit Fly	https://metazoa.ensembl.org/Drosophila_melanogaster/Gene/Summary?g=FBgn0010247;r=3R:3461351-3587054;t=FBtr0113885;db=core https://uswest.ensembl.org/Drosophila_melanogaster/Gene/Summary?g=FBgn0010247;r=3R:3461351-3587054;t=FBtr0113885 https://www.ncbi.nlm.nih.gov/gene/3355109
PARP5 Fruit Fly	https://metazoa.ensembl.org/Drosophila_melanogaster/Gene/Summary?g=FBgn0027508;r=3R:25653091-25661286;t=FBtr0084890;db=core https://uswest.ensembl.org/Drosophila_melanogaster/Gene/Summary?g=FBgn0027508;r=3R:25653091-25661286 https://www.ncbi.nlm.nih.gov/gene/43095
PARP16 Fruit Fly	https://metazoa.ensembl.org/Drosophila_melanogaster/Gene/Summary?g=FBgn0034129;r=2R:16445986-16447512;t=FBtr0087159;db=core https://www.ncbi.nlm.nih.gov/gene/36841

RESULTS

Since members of a gene family are derived from a duplication of an ancestral gene, we expect that the amount of function, sequence, expression, and genome context conservation will be related to the length of time since duplication (Andersson et al. 2015). Recently duplicated genes of a gene family will have more similar sequence, expression, and gene order than genes duplicated longer ago (Ocampo Daza et al. 2011, Tostivint et al. 2014). To investigate these lines of evidence, we have calculated amino acid identity and similarity percentages for members of each gene family, performed phylogenetic analyses for all PARPs and each species, examined pairwise-coexpression of gene family members, and investigated gene order/chromosome context.

Sequence Conservation

Functional conservation of gene family members depends upon conservation of sequences, both amino acid and DNA. Conservation of identical or similar amino acids regions within protein sequences indicate that these amino acids are important for protein function. We examined amino acid identity and similarity percentages across PARP species pairs to determine relationships amongst known grouped PARPs (Figures 2-6). We also investigated similarity of known groups to the undefined-PARPs to determine which PARP groups to which they belong (Figures 7-11).

Pairwise identity of amino acids across DNA-PARPs in human revealed some amino acid conservation amongst PARP2 and PARP3 (29.33%), PARP2 and PARP1 (21.05%), and PARP1 and PARP3 (15.56%) (Fig. 2 panel A). Tankyrase-PARPs, PARPs 5a and 5b, showed the highest conservation at 71.21%. CCCH-PARPs revealed low conservation between members PARP7 and PARP12 (13.51%), PARP7 and PARP13 (9.36%), and PARP12 and PARP13 (25.1%). In Macro-PARPs low conservation was observed amongst the pairings of PARP9 and PARP14 (12.41%) and PARP9 and PARP15 (11.37%). Undefined-PARPs 6 and 8 had moderate pairwise identity, which was the highest identified conservation amongst undefined-PARP members (44.92%).

Similarity of amino acids in the human DNA-PARPs, were similar to the identity values with PARPs 1, 2, and 3 having moderate similarity percentages with PARP1 and PARP2 (30.41%), PARP1 and PARP3 (24.78%), and PARP2 and PARP3 (42.33%) (Fig. 2 panel B). Tankyrase-PARPs showed a high similarity percentage (76.84%). CCCH-PARPs, in similar fashion, revealed amino low amino acid conservation between pairings PARP7 and PARP12 (18.87%), PARP7 and PARP13 (15.24%), and PARP12 and PARP13 (36.15%). Macro-PARP similarity percentages were also low: PARP9 and PARP14 (18.55%), PARP9 and PARP15 (17.77%), and PARP14 and PARP15 (23.63%). Undefined-PARPs with moderate pairwise similarity were PARP6 and PARP8 (53.12%), while all others were also low.

A

	PARP1_Hs	PARP2_Hs	PARP3_Hs	PARP4_Hs	PARP5a_Hs	PARP5b_Hs	PARP6_Hs	PARP7_Hs	PARP8_Hs	PARP9_Hs	PARP10_Hs	PARP11_Hs	PARP12_Hs	PARP13_Hs	PARP14_Hs	PARP15_Hs	PARP16_Hs
PARP1_Hs		21.05	15.56	5.18	7.65	7.83	6.86	3.97	9.36	3.93	3.91	3.14	4.64	5.84	4.42	5.12	4.14
PARP2_Hs	21.05		29.33	5.74	5.58	5.73	5.98	4.72	6.67	3.39	3.35	5.54	3.58	3.8	3.23	5.38	7.01
PARP3_Hs	15.56	29.33		5.15	4.37	4.97	6.25	5.6	6.09	3.94	3.1	6.53	5.08	4.77	3.66	6.3	6.3
PARP4_Hs	5.18	5.74	5.15		2.48	2.8	3.03	2.29	3.24	2.27	2.27	2.01	1.78	1.84	1.89	2.58	2.42
PARP5a_Hs	7.65	5.58	4.37	2.48		71.21	2.91	5.37	4.05	6.19	7.51	4.4	7.26	8.01	6.11	5.54	2.32
PARP5b_Hs	7.83	5.73	4.97	2.8	71.21		3.1	6.25	4.15	6.38	6.18	4.98	7.74	8.43	6.22	6.43	2.7
PARP6_Hs	6.86	5.98	6.25	3.03	2.91	3.1		4.12	44.92	2.99	2.08	3.54	5.04	4.48	2.38	2.39	8.92
PARP7_Hs	3.97	4.72	5.6	2.29	5.37	6.25	4.12		3.8	6.41	6.91	13.7	13.51	9.36	5.94	10.01	3.51
PARP8_Hs	9.36	6.67	6.09	3.24	4.05	4.15	44.92	3.8		3.2	2.8	3.13	4.23	4.59	2.7	3.32	6.57
PARP9_Hs	3.93	3.39	3.94	2.27	6.19	6.38	2.99	6.41	3.2		10.33	4.94	7.66	8.85	12.41	11.37	2.67
PARP10_Hs	3.91	3.35	3.1	2.27	7.51	6.18	2.08	6.91	2.8	10.33		5.7	7.2	6.62	8.57	9.36	2.49
PARP11_Hs	3.14	5.54	6.53	2.01	4.4	4.98	3.54	13.7	3.13	4.94	5.7		15.07	9.46	4.49	8.95	7.6
PARP12_Hs	4.64	3.58	5.08	1.78	7.26	7.74	5.04	13.51	4.23	7.66	7.2	15.07		25.1	5.84	7.23	3.81
PARP13_Hs	5.84	3.8	4.77	1.84	8.01	8.43	4.48	9.36	4.59	8.85	6.62	9.46	25.1		6.99	8.12	2.72
PARP14_Hs	4.42	3.23	3.66	1.89	6.11	6.22	2.38	5.94	2.7	12.41	8.57	4.49	5.84	6.99		19.13	1.53
PARP15_Hs	5.12	5.38	6.3	2.58	5.54	6.43	2.39	10.01	3.32	11.37	9.36	8.95	7.23	8.12	19.13		3.24
PARP16_Hs	4.14	7.01	6.3	2.42	2.32	2.7	8.92	3.51	6.57	2.67	2.49	7.6	3.81	2.72	1.53	3.24	

B

	PARP1_Hs	PARP2_Hs	PARP3_Hs	PARP4_Hs	PARP5a_Hs	PARP5b_Hs	PARP6_Hs	PARP7_Hs	PARP8_Hs	PARP9_Hs	PARP10_Hs	PARP11_Hs	PARP12_Hs	PARP13_Hs	PARP14_Hs	PARP15_Hs	PARP16_Hs
PARP1_Hs		30.41	24.78	9.77	16.12	16.2	13.2	8.33	19.06	10.38	10.54	7.46	10.29	12.32	9.26	11.43	7.82
PARP2_Hs	30.41		42.33	10.76	9.95	10.78	11.29	12.16	12.67	8.23	8.7	11.53	8.89	8.58	6.83	11.32	13.71
PARP3_Hs	24.78	42.33		9.8	8.53	9.5	12.38	12.38	12.48	6.6	7.56	12.9	9.95	9.8	7.07	12.15	14.76
PARP4_Hs	9.77	10.76	9.8		5.3	5.76	5.56	5.52	6.61	4.76	5.12	4.14	4.21	4.44	4.46	5.26	5.13
PARP5a_Hs	16.12	9.95	8.53	5.3		76.84	7.17	9.95	9.65	13.44	15.64	7.93	13.36	15.64	13.01	11.42	5.42
PARP5b_Hs	16.2	10.78	9.5	5.76	76.84		7.85	11.62	10.27	13.75	13.58	9.14	13.98	17.36	12.66	12.92	5.96
PARP6_Hs	13.2	11.29	12.38	5.56	7.17	7.85		8.25	53.12	6.32	5.44	7.62	9.07	8.71	4.81	5.34	16.43
PARP7_Hs	8.33	12.16	12.38	5.52	9.95	11.62	8.25		7.77	11.37	12.1	20.12	18.87	15.24	10.38	15.42	9.18
PARP8_Hs	19.06	12.67	12.48	6.61	9.65	10.27	53.12	7.77		6.93	7.47	6.68	8.54	9.34	6.43	7.3	12.59
PARP9_Hs	10.38	8.23	6.6	4.76	13.44	13.75	6.32	11.37	6.93		18.77	8.92	14.08	16.65	18.55	17.77	5.13
PARP10_Hs	10.54	8.7	7.56	5.12	15.64	13.58	5.44	12.1	7.47	18.77		8.51	12.44	13.61	17.1	17.62	5.15
PARP11_Hs	7.46	11.53	12.9	4.14	7.93	9.14	7.62	20.12	6.68	8.92	8.51		23.29	15.2	6.93	12.38	14.75
PARP12_Hs	10.29	8.89	9.95	4.21	13.36	13.98	9.07	18.87	8.54	14.08	12.44	23.29		36.15	10.36	10.44	8.39
PARP13_Hs	12.32	8.58	9.8	4.44	15.64	17.36	8.71	15.24	9.34	16.65	13.61	15.2	36.15		12.72	15.71	6.45
PARP14_Hs	9.26	6.83	7.07	4.46	13.01	12.66	4.81	10.38	6.43	18.55	17.1	6.93	10.36	12.72		23.63	3.58
PARP15_Hs	11.43	11.32	12.15	5.26	11.42	12.92	5.34	15.42	7.3	17.77	17.62	12.38	10.44	15.71	23.63		6.12
PARP16_Hs	7.82	13.71	14.76	5.13	5.42	5.96	16.43	9.18	12.59	5.13	5.15	14.75	8.39	6.45	3.58	6.12	

Figure. 2. Pairwise Amino Acid Identity and Similarity of Human *PARP* Proteins. 17 human protein sequences were used for pairwise amino acid identity and similarity analyses to identify pairwise percentage relationships amongst PARP members (Kent et al. 2002, Sievers et al. 2011, Stothard 2000; Table 3). Panel A in blue represents human (Hs) PARP identity and panel B in green represents human (Hs) PARP similarity.

Pairwise identity of amino acids across DNA-PARPs in mouse were similar to humans with some amino acid conservation amongst PARP2 and PARP3 (29.72%), PARP1 and PARP2 (20.82%), and PARP1 and PARP3 (14.83%) (Fig. 3 panel A). Tankyrase-PARPs, PARPs 5a and 5b, showed the highest conservation at 71.21%. CCCH-PARPs displayed low conservation between PARP7 and PARP12 (12.76%), PARP7 and PARP13 (10.19%), and PARP12 and PARP13 (21.29%). In Macro-PARPs, conservation was observed for PARP9 and PARP14 at 12.12%. Undefined-PARPs 6 and 8 had moderate pairwise identity, the highest amongst undefined-PARP members (42.81%).

Similarity among mouse DNA-PARPs also followed in similar fashion to human as DNA-PARPs 1, 2, and 3 had some low and moderate similarity between pairings PARP1 and PARP2 (30.01%), PARP1 and PARP3 (24.02%), and PARP2 and PARP3 (41.9%) (Fig. 3 panel B). Tankyrase-PARPs also showed a high percentage at 77.24%. CCCH-PARPs similarity values were: PARP7 and PARP12 (18.06%), PARP7 and PARP13 (16.78%), and PARP12 and PARP13 (31.41%). Macro-PARPs, PARP9 and PARP14 had a low similarity at 19.06%. Undefined-PARPs 6 and 8 again were the highest similarity in this group with a moderate pairwise percentage (50.55%).

A

	PARP1_Mm	PARP2_Mm	PARP3_Mm	PARP4_Mm	PARP5a_Mm	PARP5b_Mm	PARP6_Mm	PARP7_Mm	PARP8_Mm	PARP9_Mm	PARP10_Mm	PARP11_Mm	PARP12_Mm	PARP13_Mm	PARP14_Mm	PARP16_Mm
PARP1_Mm		20.82	14.83	4.69	7.76	7.83	6.69	4.05	9.23	4.14	4.18	3.17	4.94	5.51	4.31	3.96
PARP2_Mm	20.82		29.72	4.66	5.07	5.13	6.51	4.88	6.67	3.37	4.07	5.69	4.08	3.78	3.68	7.92
PARP3_Mm	14.83	29.72		4.46	4.45	4.74	6.38	5.98	6.54	4.17	3.72	7.25	5.65	4.69	3.74	6.64
PARP4_Mm	4.69	4.66	4.46		2.31	2.58	2.65	1.99	2.74	1.99	2.09	1.83	1.88	1.62	1.74	2.03
PARP5a_Mm	7.76	5.07	4.45	2.31		71.21	2.99	5.59	5.18	6.12	7.81	4.51	7.69	8.54	6.28	2.41
PARP5b_Mm	7.83	5.13	4.74	2.58	71.21		3.39	6.47	4.04	6.51	6.92	5.08	8.03	8.6	6.06	2.86
PARP6_Mm	6.69	6.51	6.38	2.65	2.99	3.39		4.02	42.81	3.15	1.7	3.71	4.89	3.93	2.42	9.21
PARP7_Mm	4.05	4.88	5.98	1.99	5.59	6.47	4.02		3.93	6.27	7.24	13.39	12.76	10.19	5.96	3.64
PARP8_Mm	9.23	6.67	6.54	2.74	5.18	4.04	42.81	3.93		3.4	2.53	3.14	4.07	4.15	2.84	6.73
PARP9_Mm	4.14	3.37	4.17	1.99	6.12	6.51	3.15	6.27	3.4		8.92	4.88	7.61	8.01	12.12	3.07
PARP10_Mm	4.18	4.07	3.72	2.09	7.81	6.92	1.7	7.24	2.53	8.92		5.76	6.44	6.39	8.58	2.45
PARP11_Mm	3.17	5.69	7.25	1.83	4.51	5.08	3.71	13.39	3.14	4.88	5.76		15	9.1	4.72	8.2
PARP12_Mm	4.94	4.08	5.65	1.88	7.69	8.03	4.89	12.76	4.07	7.61	6.44	15		21.29	6.03	4.02
PARP13_Mm	5.51	3.78	4.69	1.62	8.54	8.6	3.93	10.19	4.15	8.01	6.39	9.1	21.29		8.04	2.49
PARP14_Mm	4.31	3.68	3.74	1.74	6.28	6.06	2.42	5.96	2.84	12.12	8.58	4.72	6.03	8.04		1.88
PARP16_Mm	3.96	7.92	6.64	2.03	2.41	2.86	9.21	3.64	6.73	3.07	2.45	8.2	4.02	2.49	1.88	

B

	PARP1_Mm	PARP2_Mm	PARP3_Mm	PARP4_Mm	PARP5a_Mm	PARP5b_Mm	PARP6_Mm	PARP7_Mm	PARP8_Mm	PARP9_Mm	PARP10_Mm	PARP11_Mm	PARP12_Mm	PARP13_Mm	PARP14_Mm	PARP16_Mm
PARP1_Mm		30.01	24.02	8.81	15.95	16	13.29	8.49	18.7	10.78	11.03	7.51	9.65	11.36	9.59	7.64
PARP2_Mm	30.01		41.9	8.97	9.47	10.19	11.89	11.79	12.02	8.3	9.58	11.38	8.97	8.9	7.2	14.26
PARP3_Mm	24.02	41.9		8.55	8.37	9.26	11.82	12.8	12.13	6.9	9.04	13.69	10.05	9.89	7.48	13.62
PARP4_Mm	8.81	8.97	8.55		5.04	5.38	5	4.52	5.6	3.99	4.81	3.81	3.68	4.16	4.19	4.31
PARP5a_Mm	15.95	9.47	8.37	5.04		77.24	7.14	10.39	10.79	13.5	16.19	7.93	13.96	16.51	12.69	5.31
PARP5b_Mm	16	10.19	9.26	5.38	77.24		8	11.99	10.01	13.89	13.83	9.09	13.95	18.17	12.22	5.96
PARP6_Mm	13.29	11.89	11.82	5	7.14	8		8.15	50.55	6.98	4.74	7.42	8.98	8.33	5.12	16.43
PARP7_Mm	8.49	11.79	12.8	4.52	10.39	11.99	8.15		7.53	10.39	12.76	19.65	18.06	16.78	10.74	9.58
PARP8_Mm	18.7	12.02	12.13	5.6	10.79	10.01	50.55	7.53		7.1	6.96	6.17	7.97	9.36	6.68	12.32
PARP9_Mm	10.78	8.3	6.9	3.99	13.5	13.89	6.98	10.39	7.1		17.84	7.95	14.27	16.03	19.06	5.5
PARP10_Mm	11.03	9.58	9.04	4.81	16.19	13.83	4.74	12.76	6.96	17.84		8.83	11.45	14.01	17.05	5.26
PARP11_Mm	7.51	11.38	13.69	3.81	7.93	9.09	7.42	19.65	6.17	7.95	8.83		22.7	13.65	6.49	14.99
PARP12_Mm	9.65	8.97	10.05	3.68	13.96	13.95	8.98	18.06	7.97	14.27	11.45	22.7		31.41	10.89	8.78
PARP13_Mm	11.36	8.9	9.89	4.16	16.51	18.17	8.33	16.78	9.36	16.03	14.01	13.65	31.41		14.74	5.72
PARP14_Mm	9.59	7.2	7.48	4.19	12.69	12.22	5.12	10.74	6.68	19.06	17.05	6.49	10.89	14.74		3.76
PARP16_Mm	7.64	14.26	13.62	4.31	5.31	5.96	16.43	9.58	12.32	5.5	5.26	14.99	8.78	5.72	3.76	

Figure. 3. Pairwise Amino Acid Identity and Similarity of Mouse *PARP* Proteins. 16 mouse protein sequences were used for pairwise amino acid identity and similarity analyses to identify pairwise percentage relationships amongst PARP members (Bult et al. 2019, Sievers et al. 2011, Stothard 2000; Table 4). Panel A in blue represents mouse (Mm) PARP identity and panel B in green represents mouse (Mm) PARP similarity.

Pairwise identity of amino acids across DNA-PARPs in rat revealed some amino acid conservation amongst PARP2 and PARP3 (29.46%), PARP1 and PARP2 (20.9%), and PARP1 and PARP3 (15.37%) (Fig. 4 panel A). Tankyrase-PARPs, PARPs 5a and 5b, displayed high conservation at 71.15%. CCCH-PARPs identity showed conservation between PARP7 and PARP12 (8.63%), PARP7 and PARP13 (4.08%), and PARP12 and PARP13 (17.18%). In Macro-PARPs conservation was observed between PARP9 and PARP14 (11.73%). Undefined-PARPs 6 and 8 revealed moderate pairwise identity and were the most highly conserved pair amongst undefined-PARP members (42.45%).

Similarly, to human and mouse, rat-DNA PARPs 1, 2, and 3 had low to moderate amino acid similarity percentages with PARP1 and PARP2 paired at 29.79%, PARP1 and PARP3 at 23.54%, and PARP2 and PARP3 at 41.58% (Fig. 4 panel B). Tankyrase-PARPs showed a high similarity at 76.89%. CCCH-PARPs similarity levels were: PARP7 and PARP12 (13.1%), PARP7 and PARP13 (7.89%), and PARP12 and PARP13 (25.45%). Macro-PARPs, PARP9 and PARP14 showed low similarity (18.32%), while undefined-PARPs 6 and 8 showed a moderate pairwise similarity (50.39%).

A

	PARP1_Rn	PARP2_Rn	PARP3_Rn	PARP4_Rn	PARP5a_Rn	PARP5b_Rn	PARP6_Rn	PARP7_Rn	PARP8_Rn	PARP9_Rn	PARP10_Rn	PARP11_Rn	PARP12_Rn	PARP13_Rn	PARP14_Rn	PARP16_Rn
PARP1_Rn		20.9	15.37	4.93	7.76	7.76	6.95	4.05	9.61	4.58	4.21	2.59	3.78	3.75	4.3	4.14
PARP2_Rn	20.9		29.46	4.99	5.22	5.28	6.64	4.65	6.55	3.83	3.92	5.47	2.61	1.13	3.63	7.94
PARP3_Rn	15.37	29.46		4.79	4.47	4.83	6.17	5.75	6.16	4.47	3.75	6.46	3.59	1.42	3.64	6.54
PARP4_Rn	4.93	4.99	4.79		2.42	2.8	2.86	2.1	3.12	2.32	2.02	1.75	1.08	0.84	2.04	2.1
PARP5a_Rn	7.76	5.22	4.47	2.42		71.15	2.93	5.53	5.25	6.12	7.76	4.18	6.19	6.11	5.97	2.48
PARP5b_Rn	7.76	5.28	4.83	2.8	71.15		3.17	6.4	4.03	6.44	6.67	4.71	6.7	6.07	5.66	2.94
PARP6_Rn	6.95	6.64	6.17	2.86	2.93	3.17		3.82	42.45	3.32	1.84	3.24	3.96	2.27	2.34	9.49
PARP7_Rn	4.05	4.65	5.75	2.1	5.53	6.4	3.82		3.76	6.45	7.14	12.09	8.63	4.08	5.94	3.78
PARP8_Rn	9.61	6.55	6.16	3.12	5.25	4.03	42.45	3.76		3.39	2.64	2.67	3.2	2.63	2.95	6.6
PARP9_Rn	4.58	3.83	4.47	2.32	6.12	6.44	3.32	6.45	3.39		9.89	4.89	5.53	4.24	11.73	3.38
PARP10_Rn	4.21	3.92	3.75	2.02	7.76	6.67	1.84	7.14	2.64	9.89		5.83	4.29	4.55	8.28	2.41
PARP11_Rn	2.59	5.47	6.46	1.75	4.18	4.71	3.24	12.09	2.67	4.89	5.83		7.32	0.9	3.95	7.58
PARP12_Rn	3.78	2.61	3.59	1.08	6.19	6.7	3.96	8.63	3.2	5.53	4.29	7.32		17.18	3.98	2.61
PARP13_Rn	3.75	1.13	1.42	0.84	6.11	6.07	2.27	4.08	2.63	4.24	4.55	0.9	17.18		4.2	0.19
PARP14_Rn	4.3	3.63	3.64	2.04	5.97	5.66	2.34	5.94	2.95	11.73	8.28	3.95	3.98	4.2		1.91
PARP16_Rn	4.14	7.94	6.54	2.1	2.48	2.94	9.49	3.78	6.6	3.38	2.41	7.58	2.61	0.19	1.91	

B

	PARP1_Rn	PARP2_Rn	PARP3_Rn	PARP4_Rn	PARP5a_Rn	PARP5b_Rn	PARP6_Rn	PARP7_Rn	PARP8_Rn	PARP9_Rn	PARP10_Rn	PARP11_Rn	PARP12_Rn	PARP13_Rn	PARP14_Rn	PARP16_Rn
PARP1_Rn		29.79	23.54	9.81	15.95	16	13.55	8.56	19.06	11.44	10.67	6.21	7.64	7.64	9.53	8
PARP2_Rn	29.79		41.58	9.93	9.62	10.26	11.78	11.56	11.8	8.58	9.5	10.29	6.27	4.26	7.06	14.13
PARP3_Rn	23.54	41.58		9.38	8.6	9.43	11.39	12.22	11.93	7.76	8.66	12.76	6.74	3.99	7.81	13.59
PARP4_Rn	9.81	9.93	9.38		5.33	5.75	5.48	4.77	6.46	4.59	5.02	3.56	2.49	2.02	4.72	4.74
PARP5a_Rn	15.95	9.62	8.6	5.33		76.89	7.22	10.34	10.88	13.5	16.73	7.16	12.17	11.78	12.42	5.39
PARP5b_Rn	16	10.26	9.43	5.75	76.89		7.93	11.91	9.92	13.82	14.46	8.16	12.64	12.63	11.86	6.04
PARP6_Rn	13.55	11.78	11.39	5.48	7.22	7.93		8.15	50.39	7.16	4.79	6.06	7.72	4.95	4.84	16.43
PARP7_Rn	8.56	11.56	12.22	4.77	10.34	11.91	8.15		7.51	11.2	12.82	17.11	13.1	7.89	10.72	9.45
PARP8_Rn	19.06	11.8	11.93	6.46	10.88	9.92	50.39	7.51		7.38	6.9	5.14	6.64	5.47	6.46	12.07
PARP9_Rn	11.44	8.58	7.76	4.59	13.5	13.82	7.16	11.2	7.38		19.12	7.67	11.35	10.53	18.32	5.81
PARP10_Rn	10.67	9.5	8.66	5.02	16.73	14.46	4.79	12.82	6.9	19.12		9.14	8.43	9.55	16.26	5.36
PARP11_Rn	6.21	10.29	12.76	3.56	7.16	8.16	6.06	17.11	5.14	7.67	9.14		11.69	1.3	5.65	13.89
PARP12_Rn	7.64	6.27	6.74	2.49	12.17	12.64	7.72	13.1	6.64	11.35	8.43	11.69		25.45	8.72	6.67
PARP13_Rn	7.64	4.26	3.99	2.02	11.78	12.63	4.95	7.89	5.47	10.53	9.55	1.3	25.45		9.86	1.16
PARP14_Rn	9.53	7.06	7.81	4.72	12.42	11.86	4.84	10.72	6.46	18.32	16.26	5.65	8.72	9.86		3.76
PARP16_Rn	8	14.13	13.59	4.74	5.39	6.04	16.43	9.45	12.07	5.81	5.36	13.89	6.67	1.16	3.76	

Figure. 4. Pairwise Amino Acid Identity and Similarity of Rat *PARP* Proteins. 16 rat protein sequences were used for pairwise amino acid identity and similarity analyses to identify pairwise percentage relationships amongst PARP members (Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Table 5). Panel A in blue represents rat (Rn) PARP identity and panel B in green represents rat (Rn) PARP similarity.

Pairwise identity of amino acids across DNA-PARPs in zebrafish displayed some amino acid conservation amongst PARP2 and PARP3 (27.98%), PARP2 and PARP1 (22.72%), and PARP1 and PARP3 (16.73%) (Fig. 5 panel A). Tankyrase-PARPs, PARPs 5_1 and 5_2, had high conservation at 85.93%. CCCH-PARPs showed low conservation between PARP7 and PARP12a (2.54%) and PARP7 and PARP12b (3.2%). In Macro-PARPs low conservation is observed amongst pairings PARP9 and PARP14_1 (7.99%), PARP9 and PARP14_3 (7.48%), and PARP14_1 and PARP14_3 (31.83%). Undefined-PARPs 6a and 8 revealed low pairwise expression (9.9%), while 6b and 8 had a moderate identity of 42.32%, which was the highest identity amongst unknown members.

Zebrafish DNA-PARPs 1, 2, and 3 showed some similarity as PARP1 and PARP2 was 33.3%, PARP1 and PARP3 was 25.19%, and PARP2 and PARP3 was 38.1% (Fig. 5 panel B). Tankyrase-PARPs showed a high similarity at 89.55%. CCCH-PARPs similarity was low: PARP7 and PARP12a (4.16%) and PARP7 and PARP12b (5.95%). PARP12a and PARP12b, had a similarity of 37.67%. Macro-PARPs, PARP9 and PARP14_1 had 11.49% similarity and PARP9 and PARP14_3 had 11.55% similarity. PARP14_1 and PARP14_3 showed moderate similarity with a percentage of 49.61%. Undefined-PARPs, PARP6a and PARP8 had a similarity percentage of 14.38%, while PARP members 6b and 8 had a moderate similarity of 52.06%.

A

	PARP1_Dre	PARP2_Dre	PARP3_Dre	PARP4_Dre	PARP5_1Dre	PARP5_2Dre	PARP6a_Dre	PARP6b_Dre	PARP7_Dre	PARP8_Dre	PARP9_Dre	PARP10_Dre	PARP11_Dre	PARP12a_Dre	PARP12b_Dre	PARP14_1Dre	PARP14_2Dre	PARP14_3Dre	PARP16_Dre
PARP1_Dre		22.72	16.73	5.14	6.63	6.85	2.58	6.91	1.31	8.91	1.7	6.15	3.7	5.13	3.03	4.55	4.58	4.32	3.4
PARP2_Dre	22.72		27.98	4.65	5.35	5.33	0.75	5.96	1.85	6.4	0.67	6.74	5.6	4.31	0.83	3.78	5.45	4.35	6.24
PARP3_Dre	16.73	27.98		4.21	4.79	4.9	0	5.35	2.09	5.2	0	7.36	6.37	5.95	1.38	3.51	7.7	3.84	6.28
PARP4_Dre	5.14	4.65	4.21		2.22	2.16	0.15	2.58	0.35	3.03	0.46	2.67	2.07	2.2	0.6	2.37	2.27	2	1.94
PARP5_1Dre	6.63	5.35	4.79	2.22		85.93	0.98	3.63	1.17	4.4	2.16	5.77	3.84	5.74	2.51	5.88	5.02	5.92	2.44
PARP5_2Dre	6.85	5.33	4.9	2.16	85.93		1.21	3.7	1.26	4.48	1.78	5.82	3.8	5.59	2.76	5.37	4.99	5.75	2.17
PARP6a_Dre	2.58	0.75	0	0.15	0.98	1.21		21.24	0	9.9	0.64	0	0	0.25	0.52	0.2	0	0.15	0
PARP6b_Dre	6.91	5.96	5.35	2.58	3.63	3.7	21.24		0.11	42.32	0.19	3.66	4.25	4.41	1.31	1.7	5.07	2.13	8.29
PARP7_Dre	1.31	1.85	2.09	0.35	1.17	1.26	0	0.11		1.07	1.66	2.66	0	2.54	3.2	1.57	2.05	1.97	0
PARP8_Dre	8.91	6.4	5.2	3.03	4.4	4.48	9.9	42.32	1.07		1.01	4.35	3.07	4.16	1.69	2.83	4.43	2.98	6.62
PARP9_Dre	1.7	0.67	0	0.46	2.16	1.78	0.64	0.19	1.66	1.01		0.6	0	1.23	1.05	7.99	0	7.48	0
PARP10_Dre	6.15	6.74	7.36	2.67	5.77	5.82	0	3.66	2.66	4.35	0.6		10.26	7.53	1.69	6.77	14.84	6.82	4.14
PARP11_Dre	3.7	5.6	6.37	2.07	3.84	3.8	0	4.25	0	3.07	0	10.26		13.24	2.88	4.26	13.51	4.3	8.96
PARP12a_Dre	5.13	4.31	5.95	2.2	5.74	5.59	0.25	4.41	2.54	4.16	1.23	7.53	13.24		27.18	5.94	9.59	5.6	4.05
PARP12b_Dre	3.03	0.83	1.38	0.6	2.51	2.76	0.52	1.31	3.2	1.69	1.05	1.69	2.88	27.18		2.01	2.66	2.02	0.42
PARP14_1Dre	4.55	3.78	3.51	2.37	5.88	5.37	0.2	1.7	1.57	2.83	7.99	6.77	4.26	5.94	2.01		7.75	31.83	1.57
PARP14_2Dre	4.58	5.45	7.7	2.27	5.02	4.99	0	5.07	2.05	4.43	0	14.84	13.51	9.59	2.66	7.75		7.75	5.65
PARP14_3Dre	4.32	4.35	3.84	2	5.92	5.75	0.15	2.13	1.97	2.98	7.48	6.82	4.3	5.6	2.02	31.83	7.75		1.68
PARP16_Dre	3.4	6.24	6.28	1.94	2.44	2.17	0	8.29	0	6.62	0	4.14	8.96	4.05	0.42	1.57	5.65	1.68	

B

	PARP1_Dre	PARP2_Dre	PARP3_Dre	PARP4_Dre	PARP5_1Dre	PARP5_2Dre	PARP6a_Dre	PARP6b_Dre	PARP7_Dre	PARP8_Dre	PARP9_Dre	PARP10_Dre	PARP11_Dre	PARP12a_Dre	PARP12b_Dre	PARP14_1Dre	PARP14_2Dre	PARP14_3Dre	PARP16_Dre
PARP1_Dre		33.3	25.19	10.8	14.63	14.77	5.25	14.16	3.22	18.91	5.17	12.3	7.31	11.64	6.3	9.98	10.36	10.18	8.1
PARP2_Dre	33.3		38.1	9.96	10.3	10.46	1.62	11.91	4.57	13.56	1.73	15.64	10.5	10.02	3.01	8.43	14.07	8.6	12.62
PARP3_Dre	25.19	38.1		9.1	8.74	8.82	0	11.05	4.33	11.6	0	16.62	11.22	10.58	2.87	7.44	17.33	8.15	14.71
PARP4_Dre	10.8	9.96	9.1		5.41	5.43	0.25	4.74	1.55	5.83	1.18	5.88	4.25	4.54	1.33	5.38	4.89	5.2	4.52
PARP5_1Dre	14.63	10.3	8.74	5.41		89.55	2.87	8.33	3.95	10.59	5.67	11.41	6.78	11.62	7.17	11.67	9.46	11.67	5.02
PARP5_2Dre	14.77	10.46	8.82	5.43	89.55		2.88	8.48	4.06	10.95	5.77	11.72	6.69	11.75	7.18	11.39	9.4	11.63	5
PARP6a_Dre	5.25	1.62	0	0.25	2.87	2.88		24.5	0	14.38	1.29	0	0	1	1.55	0.56	0	0.46	0
PARP6b_Dre	14.16	11.91	11.05	4.74	8.33	8.48	24.5		0.56	52.06	0.94	8.38	8.08	8.4	3.38	4.3	9.8	4.17	16.85
PARP7_Dre	3.22	4.57	4.33	1.55	3.95	4.06	0	0.56		2.42	3.32	6.01	0	4.16	5.95	4.34	4.53	4.7	0
PARP8_Dre	18.91	13.56	11.6	5.83	10.59	10.95	14.38	52.06	2.42		2.53	10.55	6.67	8.86	4.88	6.82	8.77	6.47	13.25
PARP9_Dre	5.17	1.73	0	1.18	5.67	5.77	1.29	0.94	3.32	2.53		0.91	0	3.48	5.01	11.49	0	11.55	0
PARP10_Dre	12.3	15.64	16.62	5.88	11.41	11.72	0	8.38	6.01	10.55	0.91		19.7	13.62	4.41	11.94	26.56	12.77	9.76
PARP11_Dre	7.31	10.5	11.22	4.25	6.78	6.69	0	8.08	0	6.67	0	19.7		19.56	5.45	6.78	22.97	6.5	18.64
PARP12a_Dre	11.64	10.02	10.58	4.54	11.62	11.75	1	8.4	4.16	8.86	3.48	13.62	19.56		37.67	9.89	15.07	9.2	9.14
PARP12b_Dre	6.3	3.01	2.87	1.33	7.17	7.18	1.55	3.38	5.95	4.88	5.01	4.41	5.45	37.67		4.23	4.44	4.35	1.56
PARP14_1Dre	9.98	8.43	7.44	5.38	11.67	11.39	0.56	4.3	4.34	6.82	11.49	11.94	6.78	9.89	4.23		12.45	49.61	3.5
PARP14_2Dre	10.36	14.07	17.33	4.89	9.46	9.4	0	9.8	4.53	8.77	0	26.56	22.97	15.07	4.44	12.45		12.36	11.63
PARP14_3Dre	10.18	8.6	8.15	5.2	11.67	11.63	0.46	4.17	4.7	6.47	11.55	12.77	6.5	9.2	4.35	49.61	12.36		3.2
PARP16_Dre	8.1	12.62	14.71	4.52	5.02	5	0	16.85	0	13.25	0	9.76	18.64	9.14	1.56	3.5	11.63	3.2	

Figure. 5. Pairwise Amino Acid Identity and Similarity of Zebrafish *PARP* Proteins. 19 zebrafish protein sequences were used for pairwise amino acid identity and similarity analyses to identify pairwise percentage relationships amongst PARP members (Cunningham et al. 2022, Sievers et al. 2011, Stothard 2000; Table 6). *PARP14rs2.1* Dre member was not chosen for analysis. Panel A in blue represents zebrafish (Dre) PARP identity and panel B in green represents zebrafish (Dre) PARP similarity.

Fruit flies have only 3 PARPs and each is a different sub-group, so these were compared. Fruit fly PARPs amino acid identity was among members PARP1, PARP5, and PARP16. Identity was for PARP1 and PARP5 (6.8%), but even lower amongst the other pairs (Fig. 6 panel A). Fruit fly PARPs amino acid similarity percentages was for PARP1, PARP5, and PARP16 and were also low with PARP1 and PARP5 (14.6%) and even lower amongst other pairs (Fig. 6 panel B).

A

	PARP1_Dme	PARP5_Dme	PARP16_Dme
PARP1_Dme		6.8	2.93
PARP5_Dme	6.8		2.33
PARP16_Dme	2.93	2.33	

B

	PARP1_Dme	PARP5_Dme	PARP16_Dme
PARP1_Dme		14.6	6.96
PARP5_Dme	14.6		5.66
PARP16_Dme	6.96	5.66	

Figure. 6. Pairwise Amino Acid Identity and Similarity of Fruit Fly *PARP* Proteins. 3 fruit fly protein sequences were used for pairwise amino acid identity and similarity analyses to identify pairwise percentage relationships amongst PARP members (Cunningham et al. 2022, Yates et al. 2022, Sievers et al. 2011, Stothard 2000; Table 6). Panel A in blue represents fruit fly (Dme) PARP identity and panel B in green represents fruit fly (Dme) PARP similarity.

To gain a better idea of the closest relatives to the undefined-PARPs, each undefined-PARP was compared to the following group members: DNA, Tankyrases, CCCH, and Macro-PARPs. Undefined-PARPs similarity to DNA-PARPs across all species (human, mouse, rat, zebrafish, and fruit fly) showed low percentages, ranging from 0.00 to 19.46%, with the highest similarity being between PARP8 and PARP1 (Fig. 7). A comparison of similarity between undefined-PARPs and Tankyrase-PARPs across all species displayed low similarity ranging from 2.89 to 16.79% (Fig. 8). The highest similarity was between PARP10 and PARP5. When comparing similarity of CCCH-PARPs and undefined-PARPs across all species, amino acid similarity ranged from 0.00 to 23.56% with higher identity with PARP10 and PARP11 relative to PARP7, PARP12, and PARP13 (Fig. 9).

The comparison of Macro-PARPs to undefined-PARPs, showed similarities ranging from 0.00 to 26.56% (Fig. 10). PARP10 had the highest total similarity across all Macro-PARP members for most species. PARP11 showed some similarity to both PARP14 and PARP15, but was confined to *Danio rerio* for PARP14 (18.39-22.97%) and human for PARP15 (11.69-14.62%). Other undefined-PARPs showed very little similarity. To investigate relationships with the undefined-PARP group they were compared to each other (Fig. 11). Similarity was the highest amongst PARP6 and PARP8 proteins across species, ranging from 14.06 to 53.12%. *Danio rerio* PARP10 appears to be more similar to PARP11 regardless of species with similarity of 17.13 to 19.84%. Also, PARP11 and PARP16 also showed some identity (10.66-17.59%) across species.

	Undefined vs DNA PARPs												
	PARP1_Dme	PARP1_Dre	PARP1_Hs	PARP1_Mm	PARP1_Rn	PARP2_Dre	PARP2_Hs	PARP2_Mm	PARP2_Rn	PARP3_Dre	PARP3_Hs	PARP3_Mm	PARP3_Rn
PARP10_Dre	9.96	11.78	11.53	11.95	11.94	15.63	16.6	16.34	16.49	15.9	15.23	16.26	16.3
PARP10_Hs	10.4	10.01	10.54	10.47	10.33	8.94	8.7	8.33	8.57	8.15	7.56	8.05	8.07
PARP10_Mm	10.63	11.38	11.17	11.03	11.03	9.9	9.81	9.58	9.76	8.89	8.42	9.04	8.8
PARP10_Rn	10.68	11.01	10.67	10.67	10.67	9.87	9.29	9.16	9.5	8.82	8.44	8.73	8.66
PARP11_Dre	6.82	7.06	7.24	7.16	7.24	10.36	10.84	11.02	10.73	11.51	11.66	12.97	13.12
PARP11_Hs	6.77	7.73	7.46	7.37	7.37	10.05	11.53	11.57	11.59	11.71	12.9	13.38	14.17
PARP11_Mm	6.72	7.87	7.59	7.51	7.5	10.01	11.35	11.38	11.4	12	13.2	13.69	14.5
PARP11_Rn	5.95	6.68	6.3	6.22	6.21	8.71	10.09	10.27	10.29	10.39	11.93	11.93	12.76
PARP16_Dme	6.96	7.48	7.2	7.03	7.02	10.43	11.79	11.38	11.09	12.24	13.06	12.92	12.56
PARP16_Dre	7.76	8.1	8.36	8.28	8.27	12.62	14.77	14.87	14.58	14.71	15.4	14.76	14.57
PARP16_Hs	7.49	7.83	7.82	7.64	7.73	12.22	13.71	13.47	13.33	14.57	14.76	13.95	13.93
PARP16_Mm	7.77	7.92	7.82	7.64	7.82	12.22	14.49	14.26	14.13	14.24	14.26	13.62	13.59
PARP16_Rn	8.05	7.73	8	7.92	8	12.22	14.33	14.26	14.13	14.24	14.43	13.62	13.59
PARP4_Dre	10.14	10.8	10.83	10.66	10.75	9.96	10.46	9.9	10.01	9.1	9.05	8.88	9.01
PARP4_Hs	9.2	9.37	9.77	9.73	9.72	10.12	10.76	10.16	9.94	9.69	9.8	9.78	9.65
PARP4_Mm	8.54	8.49	8.89	8.81	8.76	9.11	9.2	8.97	8.78	8.32	8.27	8.55	8.38
PARP4_Rn	9.39	9.51	9.81	9.77	9.81	10.2	10.45	10.3	9.93	9.37	9.37	9.45	9.38
PARP6_Hs	13.29	13.36	13.2	13.29	13.64	11.56	11.29	11.89	11.78	10.52	12.38	11.82	11.39
PARP6_Mm	13.29	13.36	13.2	13.29	13.64	11.56	11.29	11.89	11.78	10.52	12.38	11.82	11.39
PARP6_Rn	13.21	13.27	13.11	13.21	13.55	11.45	11.29	11.89	11.78	10.52	12.38	11.82	11.39
PARP6a_Dre	4.66	5.25	5.16	5.26	5.54	1.62	0.13	0.13	0.13	0	0	0	0
PARP6b_Dre	14.11	14.16	14.09	13.93	14.19	11.91	11.78	12.26	12.16	11.05	12.53	12.33	12.14
PARP8_Dre	18.32	18.91	18.94	18.94	19.36	13.56	12.27	12.35	12.06	11.6	12.36	12.26	11.8
PARP8_Hs	19.46	19.28	19.06	18.97	19.39	13.37	12.67	12.66	12.56	11.82	12.48	12.38	12.21
PARP8_Mm	19.22	19.09	18.78	18.7	19.03	12.82	12.13	12.02	11.83	11.5	12.32	12.13	11.87
PARP8_Rn	19.17	18.97	18.82	18.73	19.06	12.69	12.09	11.99	11.8	11.37	12.19	12.19	11.93

Figure. 7. Pairwise Amino Acid Similarity of Undefined-PARPs relative to DNA-PARPs. Each of these DNA and undefined-PARP protein sequences from human (Hs), mouse (Mm), rat (Rn), zebrafish (Dre), and fruit fly (Dme) species were used to identify pairwise amino acid similarity across these members (Kent et al. 2002, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Tables 3-6).

	Undefined vs Tankyrase PARPs								
	PARP5_1Dre	PARP5_2Dre	PARP5_Dme	PARP5a_Hs	PARP5a_Mm	PARP5a_Rn	PARP5b_Hs	PARP5b_Mm	PARP5b_Rn
PARP10_Dre	11.71	11.88	12.29	11.12	11.17	11.2	13.36	13.59	13.52
PARP10_Hs	14.88	14.98	13.59	15.64	15.79	15.85	13.58	13.71	13.77
PARP10_Mm	15.07	15.42	13.65	15.97	16.19	16.12	13.7	13.83	13.9
PARP10_Rn	15.61	15.84	14.4	16.65	16.79	16.73	14.26	14.39	14.46
PARP11_Dre	6.93	6.84	7.27	6.7	6.73	6.74	7.75	7.83	7.83
PARP11_Hs	7.99	8.06	8.78	7.93	7.97	7.99	9.14	9.22	9.22
PARP11_Mm	7.87	7.94	8.7	7.89	7.93	7.94	9.01	9.09	9.09
PARP11_Rn	7.14	7.29	7.93	7.11	7.15	7.16	8.07	8.16	8.16
PARP16_Dme	4.96	4.93	5.66	4.8	4.75	4.83	5.48	5.48	5.4
PARP16_Dre	5.52	5.49	6.03	5.19	5.15	5.23	5.85	5.78	5.78
PARP16_Hs	5.54	5.59	5.88	5.42	5.38	5.46	5.96	5.96	5.96
PARP16_Mm	5.46	5.51	6.04	5.35	5.31	5.39	5.96	5.96	5.96
PARP16_Rn	5.46	5.51	6.04	5.35	5.31	5.39	6.04	6.04	6.04
PARP4_Dre	5.7	5.69	5.95	5.64	5.65	5.66	5.78	5.7	5.74
PARP4_Hs	5.36	5.47	5.69	5.3	5.32	5.32	5.76	5.76	5.76
PARP4_Mm	5.07	5.1	5.29	5.02	5.04	5.04	5.46	5.38	5.38
PARP4_Rn	5.36	5.43	5.69	5.31	5.32	5.33	5.83	5.75	5.75
PARP6_Hs	7.59	7.58	8.48	7.17	7.14	7.22	7.85	8	7.93
PARP6_Mm	7.59	7.58	8.48	7.17	7.14	7.22	7.85	8	7.93
PARP6_Rn	7.59	7.58	8.48	7.17	7.14	7.22	7.85	8	7.93
PARP6a_Dre	2.99	3.02	3.06	2.89	2.9	2.91	3.27	3.19	3.27
PARP6b_Dre	7.95	7.94	9	7.52	7.49	7.57	8.38	8.52	8.45
PARP8_Dre	9.73	9.88	10.69	9.69	9.6	9.68	10.24	10.37	10.31
PARP8_Hs	9.63	9.9	10.38	9.65	9.49	9.57	10.27	10.4	10.33
PARP8_Mm	9.97	10.69	10.05	11.01	10.79	10.88	9.87	10.01	9.94
PARP8_Rn	10.03	10.69	10.03	11.01	10.79	10.88	9.85	9.99	9.92

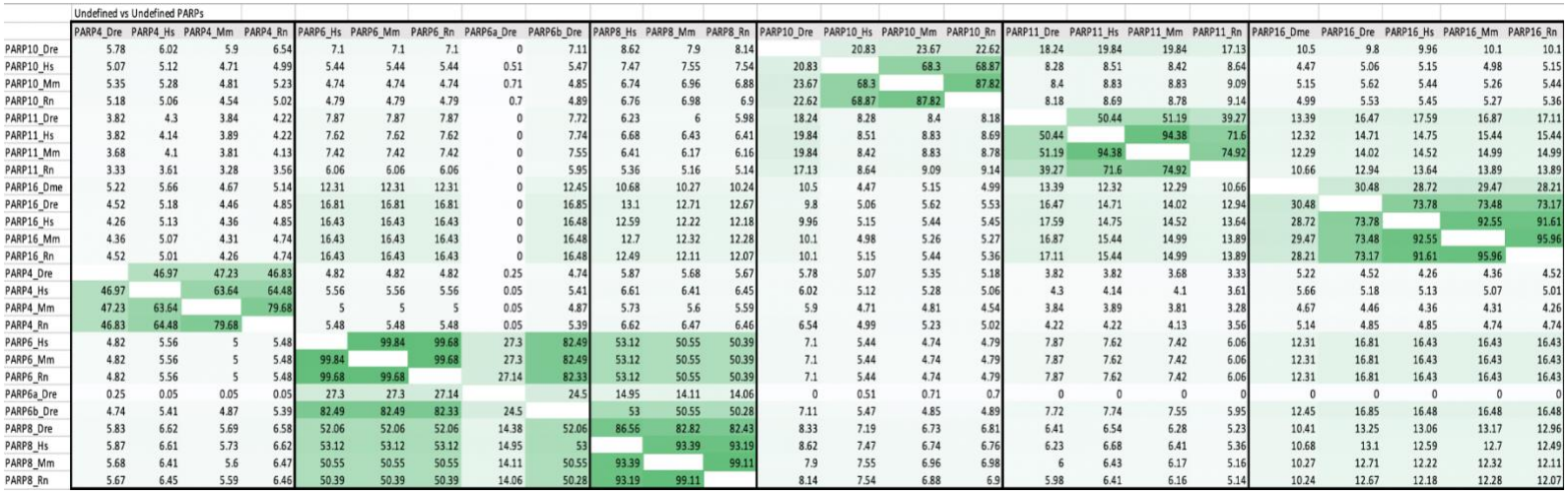
Figure. 8. Pairwise Amino Acid Similarity of Undefined-PARPs relative to Tankyrase-PARPs. Each of these Tankyrase and undefined-PARP protein sequences from human (Hs), mouse (Mm), rat (Rn), zebrafish (Dre), and fruit fly (Dme) species were used to identify pairwise amino acid similarity across these members (Kent et al. 2002, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Tables 3-6).

	Undefined vs CCCH PARPs											
	PARP7_Dre	PARP7_Hs	PARP7_Mm	PARP7_Rn	PARP12_Hs	PARP12_Mm	PARP12_Rn	PARP12a_Dre	PARP12b_Dre	PARP13_Hs	PARP13_Mm	PARP13_Rn
PARP10_Dre	5.57	20.02	19.78	19.78	14.73	13.93	10.11	14.51	4.44	12.96	13.26	6.76
PARP10_Hs	4.85	12.1	12.25	12.18	12.44	12.49	8.71	11.28	5	13.61	14.37	9.35
PARP10_Mm	5.22	12.67	12.76	12.67	11.57	11.45	8.16	10.61	4.49	12.52	14.01	8.92
PARP10_Rn	5.32	12.74	12.91	12.82	11.69	11.64	8.43	11.45	4.92	13.02	14.71	9.55
PARP11_Dre	0	17.39	17.22	17.07	19.78	18.67	13.67	19.05	5.34	13.2	11.81	3.65
PARP11_Hs	0	20.12	20.09	19.94	23.29	22.7	16.34	22.94	6.57	15.2	13.93	4.3
PARP11_Mm	0	19.68	19.65	19.51	23.56	22.7	16.34	23.41	6.43	14.79	13.65	4.2
PARP11_Rn	0	17.13	17.11	17.11	18.77	18.38	11.69	18.06	1.35	11.7	10.84	1.3
PARP16_Dme	0.98	8.67	9.07	9.07	8.85	9.11	7.9	8.29	3.08	7.38	6.33	2.31
PARP16_Dre	0.17	9.45	9.45	9.45	9.28	8.78	6.92	9.56	1.7	6.65	5.91	1.16
PARP16_Hs	0.17	9.18	9.31	9.31	8.39	8.41	6.41	8.51	1.85	6.45	5.54	1.16
PARP16_Mm	0.17	9.58	9.58	9.58	8.77	8.78	6.8	8.91	2.13	6.55	5.72	1.26
PARP16_Rn	0.17	9.45	9.45	9.45	8.77	8.53	6.67	9.04	1.99	6.55	5.63	1.16
PARP4_Dre	1.6	5.09	4.95	5	3.74	3.6	2.19	3.9	1.1	4.21	4.18	1.63
PARP4_Hs	1.86	5.52	5.37	5.32	4.21	4.19	2.72	4.05	1.06	4.44	4.74	2.09
PARP4_Mm	1.27	4.6	4.52	4.47	3.73	3.68	2.44	3.75	0.9	4.12	4.16	1.72
PARP4_Rn	1.43	4.77	4.73	4.77	3.78	3.72	2.49	3.89	1.02	4.62	4.66	2.02
PARP6_Hs	1.86	8.25	8.15	8.15	9.07	8.98	7.61	9.13	4.89	8.71	8.33	4.95
PARP6_Mm	1.86	8.25	8.15	8.15	9.07	8.98	7.61	9.13	4.89	8.71	8.33	4.95
PARP6_Rn	1.86	8.25	8.15	8.15	9.17	9.08	7.72	9.23	5	8.71	8.33	4.95
PARP6a_Dre	0.42	0	0	0	1.96	1.94	2.06	1.92	3.56	2.07	2.16	2.7
PARP6b_Dre	2.07	8.95	8.85	8.85	8.86	8.97	7.62	9.02	4.82	8.7	8.54	5.05
PARP8_Dre	1.8	7.65	7.65	7.65	8.68	8.35	7.11	8.42	4.45	9	9.43	5.69
PARP8_Hs	1.8	7.77	7.77	7.77	8.54	8.22	6.88	8.54	4.56	9.34	9.83	5.49
PARP8_Mm	1.83	7.61	7.53	7.53	8.29	7.97	6.66	8.19	4.32	9.01	9.36	5.49
PARP8_Rn	1.83	7.59	7.51	7.51	8.27	7.95	6.64	8.17	4.31	8.92	9.34	5.47

Figure. 9. Pairwise Amino Acid Similarity of Undefined-PARPs relative to CCCH-PARPs. Each of these CCCH and undefined-PARP protein sequences from human (Hs), mouse (Mm), rat (Rn), zebrafish (Dre), and fruit fly (Dme) species were used to identify pairwise amino acid similarity across these members (Kent et al. 2002, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Tables 3-6).

	Undefined vs Macro PARPs										
	PARP9_Dre	PARP9_Hs	PARP9_Mm	PARP9_Rn	PARP14_1Dre	PARP14_2Dre	PARP14_3Dre	PARP14_Hs	PARP14_Mm	PARP14_Rn	PARP15_Hs
PARP10_Dre	0.91	12.93	12.36	12.45	11.94	26.56	12.77	12.43	12.27	12.68	18.47
PARP10_Hs	9.11	17.31	17.9	18.15	14.75	11.17	16.54	15.39	15.59	15.39	19.37
PARP10_Mm	9.07	17.62	18.26	18.66	15.27	12.19	15.89	15.06	15.27	15.02	17.45
PARP10_Rn	9.14	18.06	18.7	19.02	15.68	12.01	15.8	15.43	15.63	15.28	18.16
PARP11_Dre	0	8.31	7.98	8.2	6.78	22.97	6.5	6.59	6.59	6.86	11.69
PARP11_Hs	0	9.34	7.82	7.93	7.54	21.52	7.74	7.39	6.95	7.34	13.01
PARP11_Mm	0	8.79	7.4	7.5	7.43	21.87	7.59	7.19	6.85	7.19	13.12
PARP11_Rn	0	7.9	7.13	7.25	6.13	18.39	6.34	5.94	5.67	5.93	14.62
PARP16_Dme	0	6.25	5.67	5.88	3.94	11.41	3.7	3.33	3.04	3.41	5.21
PARP16_Dre	0	4.68	5.67	5.67	3.5	11.63	3.2	3.4	3.43	3.38	5.69
PARP16_Hs	0	4.92	5.7	5.6	3.87	12.92	3.48	3.57	3.75	3.76	6.32
PARP16_Mm	0	5.24	5.81	5.6	3.98	13.59	3.69	3.84	3.96	4.02	6.21
PARP16_Rn	0	5.35	6.12	5.91	3.93	13.09	3.64	3.78	3.96	4.02	6.21
PARP4_Dre	1.18	4.46	4.11	4.2	5.38	4.89	5.2	5.41	5.13	5.18	5.92
PARP4_Hs	1.09	4.61	4.14	4.64	5.06	5.28	5.1	4.97	5.01	5.12	5.37
PARP4_Mm	0.93	3.94	3.49	3.81	4.5	4.9	4.63	4.53	4.36	4.61	4.83
PARP4_Rn	1	4.36	4	4.3	4.85	5.31	4.99	4.91	4.73	4.97	5.32
PARP6_Hs	0.95	5.25	5.55	5.88	4.28	9.64	4.15	4.03	3.95	4.1	4.77
PARP6_Mm	0.95	5.25	5.55	5.88	4.28	9.64	4.15	4.03	3.95	4.1	4.77
PARP6_Rn	0.95	5.25	5.55	5.88	4.28	9.64	4.15	4.03	3.95	4.1	4.77
PARP6a_Dre	1.29	0.91	0.6	0.7	0.56	0	0.46	0.51	0.35	0.31	0.7
PARP6b_Dre	0.94	5.19	5.66	5.9	4.3	9.8	4.17	4.39	4.21	4.31	5.33
PARP8_Dre	2.53	6.79	7.6	7.83	6.82	8.77	6.47	6.43	6.66	6.79	7.1
PARP8_Hs	2.54	6.96	7.85	7.78	6.64	8.62	6.57	6.58	6.58	6.7	6.71
PARP8_Mm	3.37	7.5	8.23	8.3	6.89	8.33	6.69	6.69	6.75	6.97	6.35
PARP8_Rn	3.36	7.55	8.36	8.51	6.89	8.3	6.73	6.64	6.69	6.91	6.33

Figure. 10. Pairwise Amino Acid Similarity of Undefined-PARPs relative to Macro-PARPs. Each of these Macro and undefined-PARP protein sequences from human (Hs), mouse (Mm), rat (Rn), zebrafish (Dre), and fruit fly (Dme) species were used to identify pairwise amino acid similarity across these members (Kent et al. 2002, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Tables 3-6). *PARP14rs2.1* Dre member was not chosen for analysis.



Undefined vs Undefined PARPs																								
	PARP4_Dre	PARP4_Hs	PARP4_Mm	PARP4_Rn	PARP6_Hs	PARP6_Mm	PARP6_Rn	PARP6a_Dre	PARP6b_Dre	PARP8_Hs	PARP8_Mm	PARP8_Rn	PARP10_Dre	PARP10_Hs	PARP10_Mm	PARP10_Rn	PARP11_Dre	PARP11_Hs	PARP11_Mm	PARP11_Rn	PARP16_Dme	PARP16_Dre	PARP16_Hs	PARP16_Mm	PARP16_Rn
PARP10_Dre	5.78	6.02	5.9	6.54	7.1	7.1	7.1	0	7.11	8.62	7.9	8.14		20.83	23.67	22.62	18.24	19.84	19.84	17.13	10.5	9.8	9.96	10.1	10.1
PARP10_Hs	5.07	5.12	4.71	4.99	5.44	5.44	5.44	0.51	5.47	7.47	7.55	7.54	20.83		68.3	68.87	8.28	8.51	8.42	8.64	4.47	5.06	5.15	4.98	5.15
PARP10_Mm	5.35	5.28	4.81	5.23	4.74	4.74	4.74	0.71	4.85	6.74	6.96	6.88	23.67	68.3		87.82	8.4	8.83	8.83	9.09	5.15	5.62	5.44	5.26	5.44
PARP10_Rn	5.18	5.06	4.54	5.02	4.79	4.79	4.79	0.7	4.89	6.76	6.98	6.9	22.62	68.87	87.82		8.18	8.69	8.78	9.14	4.99	5.53	5.45	5.27	5.36
PARP11_Dre	3.82	4.3	3.84	4.22	7.87	7.87	7.87	0	7.72	6.23	6	5.98	18.24	8.28	8.4	8.18		50.44	51.19	39.27	13.39	16.47	17.59	16.87	17.11
PARP11_Hs	3.82	4.14	3.89	4.22	7.62	7.62	7.62	0	7.74	6.68	6.43	6.41	19.84	8.51	8.83	8.69	50.44		94.38	71.6	12.32	14.71	14.75	15.44	15.44
PARP11_Mm	3.68	4.1	3.81	4.13	7.42	7.42	7.42	0	7.55	6.41	6.17	6.16	19.84	8.42	8.83	8.78	51.19	94.38		74.92	12.29	14.02	14.52	14.99	14.99
PARP11_Rn	3.33	3.61	3.28	3.56	6.06	6.06	6.06	0	5.95	5.36	5.16	5.14	17.13	8.64	9.09	9.14	39.27	71.6	74.92		10.66	12.94	13.64	13.89	13.89
PARP16_Dme	5.22	5.66	4.67	5.14	12.31	12.31	12.31	0	12.45	10.68	10.27	10.24	10.5	4.47	5.15	4.99	13.39	12.32	12.29	10.66		30.48	28.72	29.47	28.21
PARP16_Dre	4.52	5.18	4.46	4.85	16.81	16.81	16.81	0	16.85	13.1	12.71	12.67	9.8	5.06	5.62	5.53	16.47	14.71	14.02	12.94	30.48		73.78	73.48	73.17
PARP16_Hs	4.26	5.13	4.36	4.85	16.43	16.43	16.43	0	16.48	12.59	12.22	12.18	9.96	5.15	5.44	5.45	17.59	14.75	14.52	13.64	28.72	73.78		92.55	91.61
PARP16_Mm	4.36	5.07	4.31	4.74	16.43	16.43	16.43	0	16.48	12.7	12.32	12.28	10.1	4.98	5.26	5.27	16.87	15.44	14.99	13.89	29.47	73.48	92.55		95.96
PARP16_Rn	4.52	5.01	4.26	4.74	16.43	16.43	16.43	0	16.48	12.49	12.11	12.07	10.1	5.15	5.44	5.36	17.11	15.44	14.99	13.89	28.21	73.17	91.61	95.96	
PARP4_Dre		46.97	47.23	46.83	4.82	4.82	4.82	0.25	4.74	5.87	5.68	5.67	5.78	5.07	5.35	5.18	3.82	3.82	3.68	3.33	5.22	4.52	4.26	4.36	4.52
PARP4_Hs	46.97		63.64	64.48	5.56	5.56	5.56	0.05	5.41	6.61	6.41	6.45	6.02	5.12	5.28	5.06	4.3	4.14	4.1	3.61	5.66	5.18	5.13	5.07	5.01
PARP4_Mm	47.23	63.64		79.68	5	5	5	0.05	4.87	5.73	5.6	5.59	5.9	4.71	4.81	4.54	3.84	3.89	3.81	3.28	4.67	4.46	4.36	4.31	4.26
PARP4_Rn	46.83	64.48	79.68		5.48	5.48	5.48	0.05	5.39	6.62	6.47	6.46	6.54	4.99	5.23	5.02	4.22	4.22	4.13	3.56	5.14	4.85	4.85	4.74	4.74
PARP6_Hs	4.82	5.56	5	5.48		99.84	99.68	27.3	82.49	53.12	50.55	50.39	7.1	5.44	4.74	4.79	7.87	7.62	7.42	6.06	12.31	16.81	16.43	16.43	16.43
PARP6_Mm	4.82	5.56	5	5.48	99.84		99.68	27.3	82.49	53.12	50.55	50.39	7.1	5.44	4.74	4.79	7.87	7.62	7.42	6.06	12.31	16.81	16.43	16.43	16.43
PARP6_Rn	4.82	5.56	5	5.48	99.68	99.68		27.14	82.33	53.12	50.55	50.39	7.1	5.44	4.74	4.79	7.87	7.62	7.42	6.06	12.31	16.81	16.43	16.43	16.43
PARP6a_Dre	0.25	0.05	0.05	0.05	27.3	27.3	27.14		24.5	14.95	14.11	14.06	0	0.51	0.71	0.7	0	0	0	0	0	0	0	0	0
PARP6b_Dre	4.74	5.41	4.87	5.39	82.49	82.49	82.33	24.5		53	50.55	50.28	7.11	5.47	4.85	4.89	7.72	7.74	7.55	5.95	12.45	16.85	16.48	16.48	16.48
PARP8_Dre	5.83	6.62	5.69	6.58	52.06	52.06	52.06	14.38	52.06	86.56	82.82	82.43	8.33	7.19	6.73	6.81	6.41	6.54	6.28	5.23	10.41	13.25	13.06	13.17	12.96
PARP8_Hs	5.87	6.61	5.73	6.62	53.12	53.12	53.12	14.95	53		93.39	93.19	8.62	7.47	6.74	6.76	6.23	6.68	6.41	5.36	10.68	13.1	12.59	12.7	12.49
PARP8_Mm	5.68	6.41	5.6	6.47	50.55	50.55	50.55	14.11	50.55	93.39		99.11	7.9	7.55	6.96	6.98	6	6.43	6.17	5.16	10.27	12.71	12.22	12.32	12.11
PARP8_Rn	5.67	6.45	5.59	6.46	50.39	50.39	50.39	14.06	50.28	93.19	99.11		8.14	7.54	6.88	6.9	5.98	6.41	6.16	5.14	10.24	12.67	12.18	12.28	12.07

Figure. 11. Pairwise Amino Acid Similarity of Undefined-PARPs relative to Undefined-PARPs. Each of these undefined-PARP protein sequences from human (Hs), mouse (Mm), rat (Rn) , zebrafish (Dre), and fruit fly (Dme) species were used to identify pairwise amino acid similarity across these members (Kent et al. 2002, Cunningham et al. 2022, Yates et al. 2022, Bult et al. 2019, Smith et al. 2019, Sievers et al. 2011, Stothard 2000; Tables 3-6).

In order to continue analysis of amino acid conservation and divergence with the *PARP* gene family, we performed visual multiple alignments using the COBALT multiple alignment tool for five model organism species (Papadopoulos and Agarwala 2007, NCBI 2022; Tables 3-6) to identify the regions of conserved amino acids, particularly those that are likely important for PARP function (Figures 12-16).

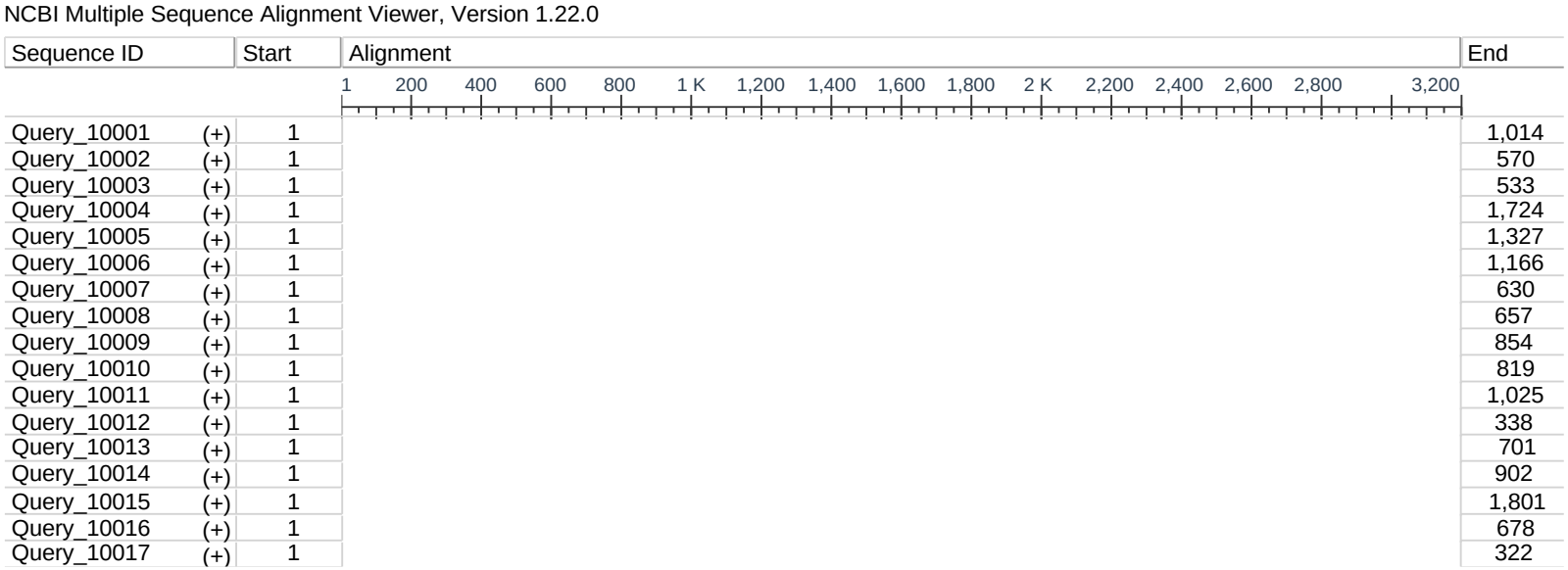


Figure. 12. COBALT-Style Multiple Sequence Alignment of Human *PARP* Proteins. 17 human PARP protein sequences were obtained from UCSC and input into the NCBI COBALT (constraint-based) multiple alignment tool (Kent et al. 2002, NCBI 2022, Papadopoulos and Agarwala 2007; Table 3). The red thin colored lines indicate some conservation of similar amino acids, and the blue thin colored lines indicates low conservation of similar amino acids. The sequence ID numbers tab that contains Query_10001 to Query_10017 represents PARP members 1, 2, 3, 4, 5a, 5b, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, and 16. The end tab represents the identified number of amino acids in each of the sequences.

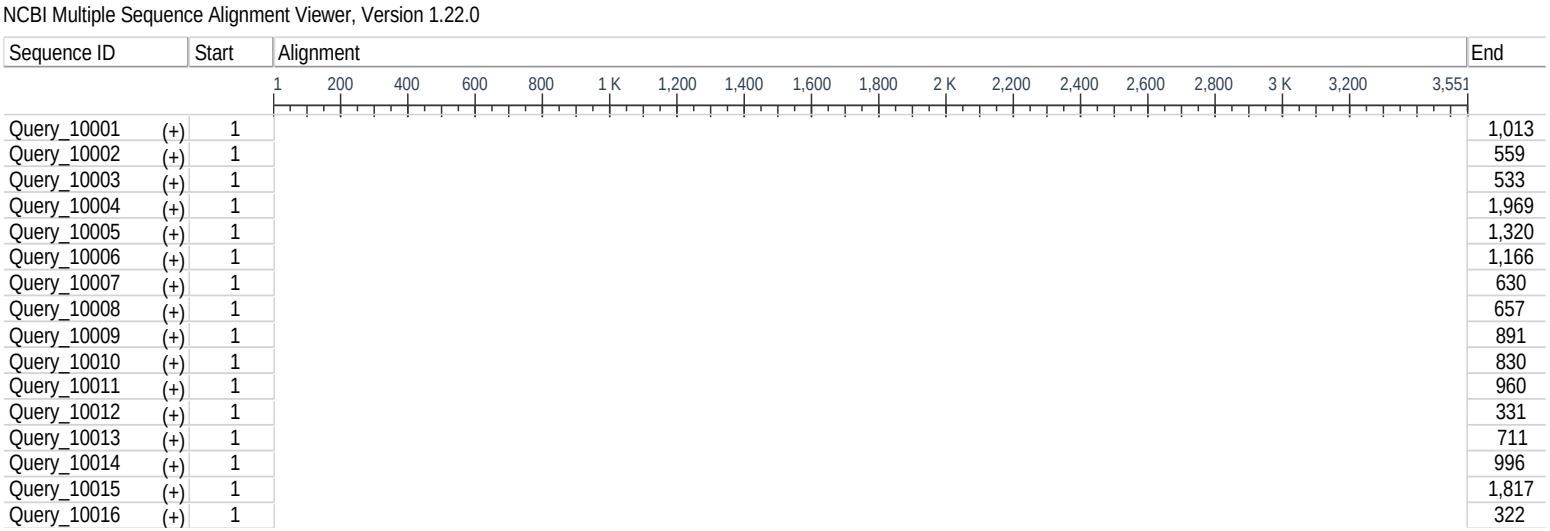


Figure. 13. COBALT-Style Multiple Sequence Alignment of Mouse *PARP* Proteins. 16 mouse PARP protein sequences were obtained from MGI and input into the NCBI COBALT (constraint-based) multiple alignment tool (Bult et al. 2019, NCBI 2022, Papadopoulos and Agarwala 2007; Table 4). The red thin colored lines indicate some conservation of similar amino acids, and the blue thin colored lines indicates low conservation of similar amino acids. The sequence ID numbers tab that contains Query_10001 to Query_10016 represents PARP members 1, 2, 3, 4, 5a, 5b, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 16. The end tab represents the identified number of amino acids in each of the sequences.

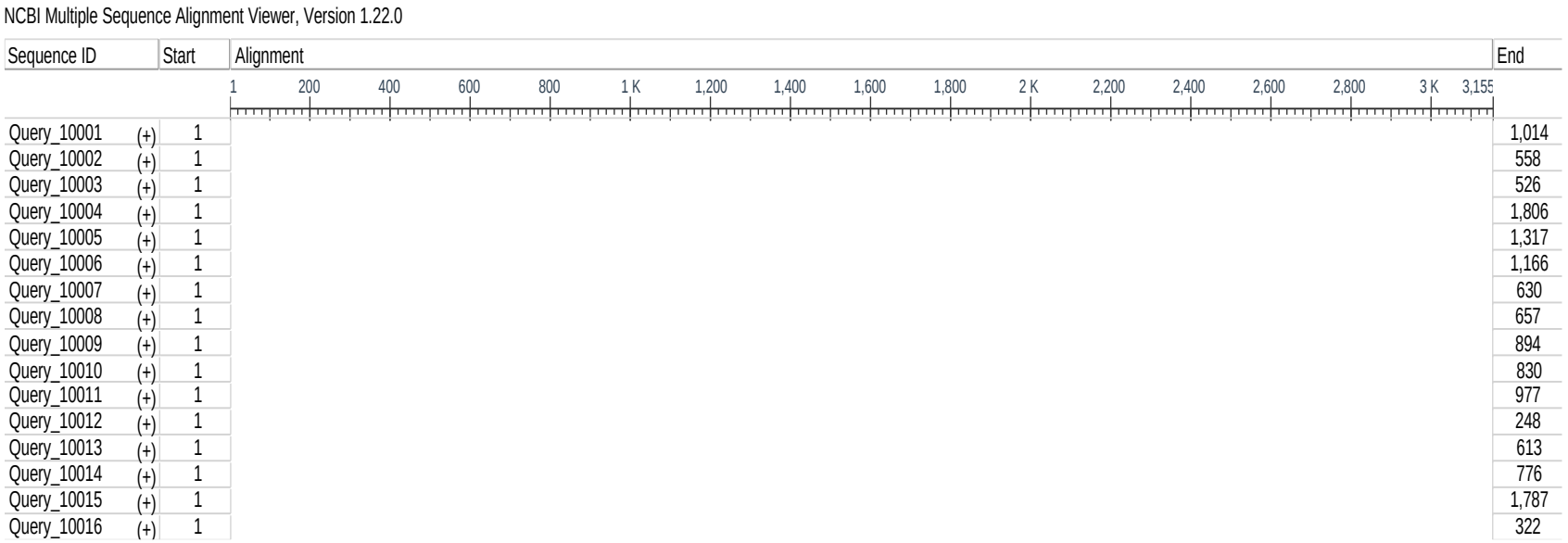


Figure. 14. COBALT-Style Multiple Sequence Alignment of Rat *PARP* Proteins. 16 rat PARP protein sequences were obtained from RGD and input into the NCBI COBALT (constraint-based) multiple alignment tool (Smith et al. 2019, NCBI 2022, Papadopoulos and Agarwala 2007; Table 5). The red thin colored lines indicate some conservation of similar amino acids, and the blue thin colored lines indicate low conservation of similar amino acids. The sequence ID numbers tab that contains Query_10001 to Query_10016 represents PARP members 1, 2, 3, 4, 5a, 5b, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 16. The end tab represents the identified number of amino acids in each of the sequences.

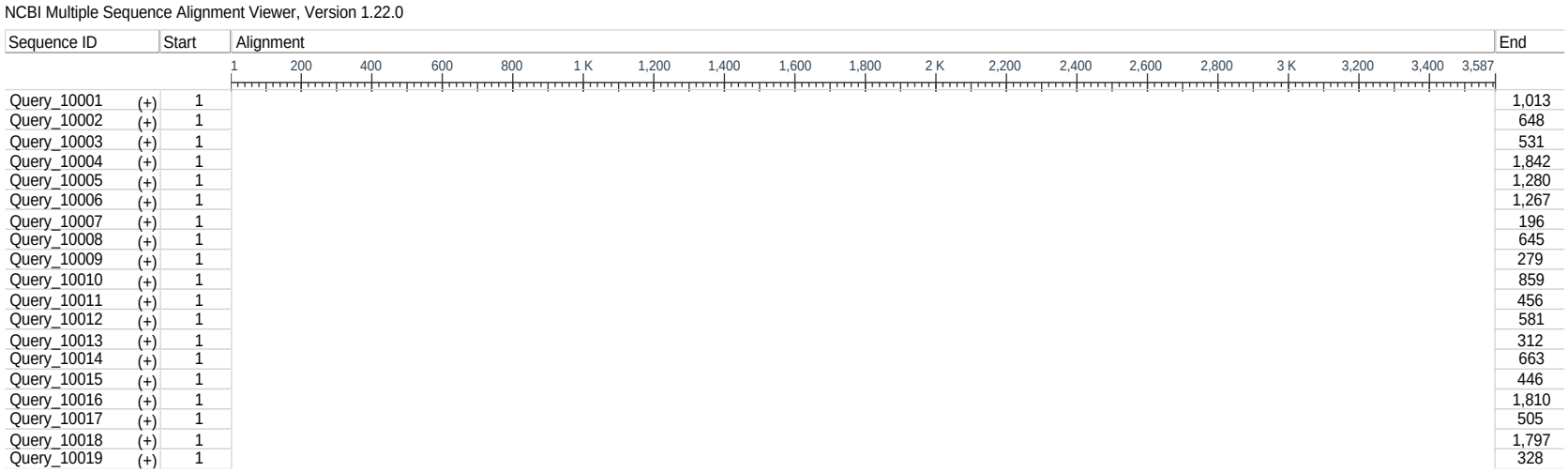


Figure. 15. COBALT-Style Multiple Sequence Alignment of Zebrafish *PARP* Proteins. 19 zebrafish PARP protein sequences were obtained from Ensembl Main and input into the NCBI COBALT (constraint-based) multiple alignment tool (Cunningham et al. 2022, NCBI 2022, Papadopoulos and Agarwala 2007; Table 6). The red thin colored lines indicate some conservation of similar amino acids, and the blue thin colored lines indicates low conservation of similar amino acids. The sequence ID numbers tab that contains Query_10001 to Query_10019 represents PARP members 1, 2, 3, 4, 5_1, 5_2, 6a, 6b, 7, 8, 9, 10, 11, 12a, 12b, 14_1, 14_2, 14_3, and 16. *PARP14rs2.1* member was not chosen for analysis. The end tab represents the identified number of amino acids in each of the sequence.

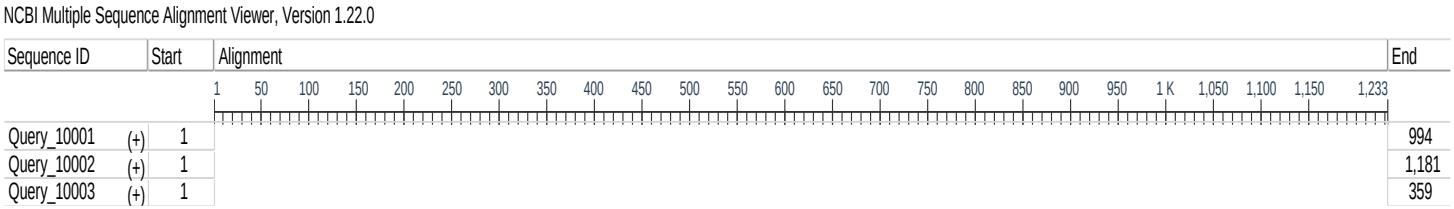


Figure. 16. COBALT-Style Multiple Sequence Alignment of Fruit Fly *PARP* Proteins. 3 fruit fly PARP protein sequences were obtained from Ensembl Main and Metazoan Ensembl and input into the NCBI COBALT (constraint-based) multiple alignment tool (Cunningham et al. 2022, Yates et al. 2022, NCBI 2022, Papadopoulos and Agarwala 2007; Table 6). The red thick colored boxes indicate high conservation of similar amino acids, the red thin colored lines indicate some conservation of similar amino acids, and the blue thin colored lines indicate low conservation of similar amino acids. The sequence ID numbers tab that contains Query_10001 to Query_10003 represents PARP members 1, 5, and 16. The end tab represents the identified number of amino acids in each of the sequences.

The COBALT multiple alignment of human PARP proteins revealed some conservation of similar amino acids (red) or low conservation (blue) across members at region 1,500 to 1,800 of the alignment with the exception of Macro-PARPs, PARP9 and PARP14 (Fig. 12). Mouse PARP protein members 6 and 7, and 9 to 13 showed some amino acid conservation around region 1,550 of the alignment (Fig. 13). Rat PARP protein members 12 and 13 (CCCH-PARPs) show amino acid conservation at region 1,000 to 1,600, while PARP members 5b and 10 from region 300-430 display some conservation (Fig. 14). Zebrafish PARP members 1-6b, 8, 10, and 14_2 , have some amino acid conservation at region 1,680 of the alignment, while members 1-5_2, 10, 11, and 14_2 have conservation at region 2,140 (Fig. 15). PARP members 10 and 11 at regions 1,980, 2,140, and 2,300 and members 14_1 and 14_3 from regions 900 to 1,400 show some amino acid conservation. All fruit fly PARP protein members 1, 5, and 16 showed high conservation of similar amino acids (red) relative to region 370-780 of the alignment (Fig. 16).

Phylogenetic Trees

While sequence alignment can provide conservation information across protein sequences, another method that provides greater detail, confidence measures, and visualization of relationships is phylogenetic analysis. Phylogenetic analysis mostly grouped the human PARP family members consistent with their functional categorization, e.g., DNA-PARPs, Tankyrase-PARPs, CCCH-PARPs, Macro-PARPs, with undefined-PARPs being dispersed (Fig. 17). Analysis confirms PARP4, PARP16, PARP6,

and PARP8, which are all undefined-PARPs, are related to one another and appear to be more closely related with the DNA-PARPs. Undefined-PARP11 grouped with the CCCH-PARPs and PARP7 appears to be closely related to PARP11 and other CCCH-PARPs. Undefined-PARP10 appears to be distantly grouped with Macro, CCCH, and Tankyrase-PARPs. For the mouse PARP comparison similarly, two large branches were observed but with more variation among functional groupings (Fig. 18). Similarly, to what was observed in the human, mouse DNA-PARPs appear to be more closely related to most undefined-PARPs and undefined-PARP11 groups with the CCCH-PARPs. Phylogenetic analysis of rat PARP members showed several branches with low confidence (Fig. 19). There is high confidence in the close relationship of DNA-PARPs and Tankyrase-PARPs, as well as undefined-PARP11 with the CCCH-PARP7. The zebrafish PARP members phylogenetic analysis showed two large groups with low confidence and one member being very divergent (PARP16) (Fig. 20). Most functional groups were found to be closely related with high confidence (Tankyrases, DNA, and Macro-PARPs). Most undefined-PARPs are found on one of the two large branches. Undefined-PARPs are grouped together with the exception of PARP4 and PARP10. Undefined-PARP10 appeared to be most closely related to the Macro-PARPs.

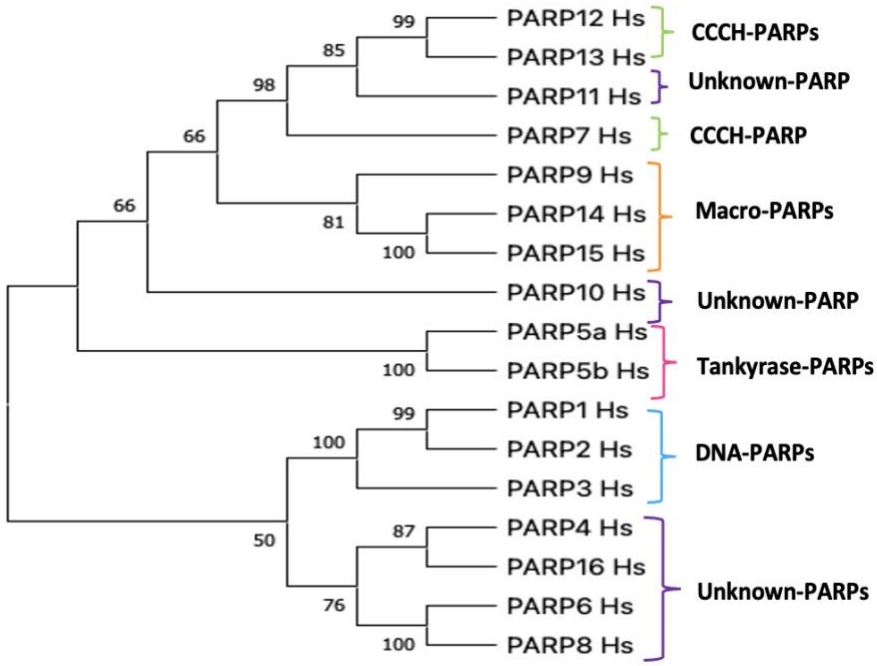


Figure. 17. Clustering by Amino Acid Similarity Shows Human *PARP* Genes Primarily Group by Function. The evolutionary history was determined using the Neighbor-Joining tree method (Saitou and Nei 1987). The percentage of replicate trees in which the associated clade clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965). All ambiguous positions were removed for each sequence pair (pairwise deletion option). A total of 2122 positions were found in the final dataset and this analysis involved 17 PARP amino acid sequences from the UCSC Genome Browser (Kent et al. 2002). Each of the PARP groups are labeled and colored according to function, green indicates the CCCH-PARPs, purple indicates the unknown (undefined)-PARPs, orange indicates the Macro-PARPs, pink indicates the Tankyrase-PARPs, and blue indicates the DNA-PARPs.

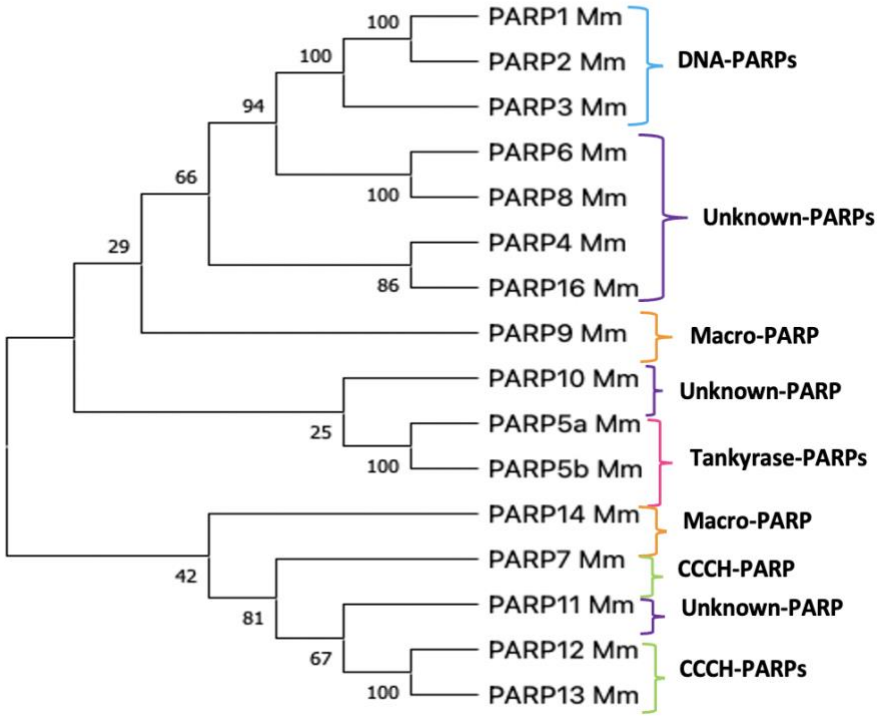


Figure. 18. Clustering by Amino Acid Similarity Shows Mouse *PARP* Genes Primarily Group by Function. The evolutionary history was determined using the Neighbor-Joining tree method (Saitou and Nei 1987). The percentage of replicate trees in which the associated clade clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965). All ambiguous positions were removed for each sequence pair (pairwise deletion option). A total of 2072 positions were found in the final dataset. This analysis involved 16 PARP amino acid sequences from the Mouse Genome Informatics Database (Bult et al. 2019).

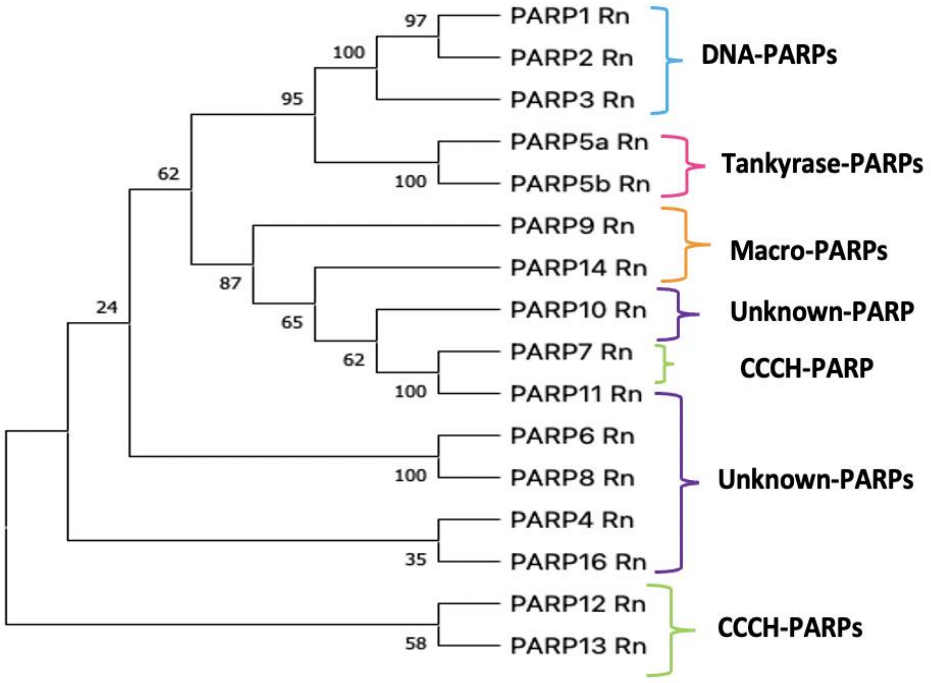


Figure. 19. Clustering by Amino Acid Similarity Shows Rat *PARP* Genes Primarily Group by Function. The evolutionary history was determined using the Neighbor-Joining tree method (Saitou and Nei 1987). The percentage of replicate trees in which the associated clade clustered together in the bootstrap test (499 replicates) are shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965). All ambiguous positions were removed for each sequence pair (pairwise deletion option). A total of 1928 positions were found in the final dataset and this analysis involved 16 PARP amino acid sequences from the Rat Genome Database (Smith et al. 2019).

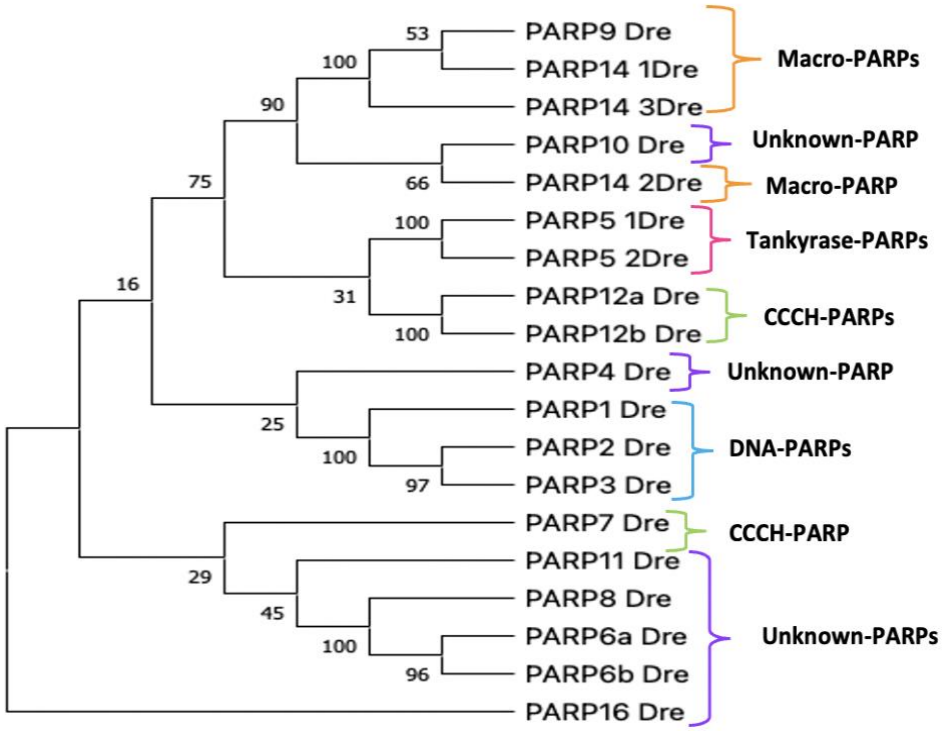


Figure. 20. Clustering by Amino Acid Similarity Shows Zebrafish *PARP* Genes Group Primarily by Function. The evolutionary history was determined using the Neighbor-Joining tree method (Saitou and Nei 1987). The percentage of replicate trees in which the associated clade clustered together in the bootstrap test (489 replicates) are shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965). All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 2032 positions in the final dataset. This analysis involved 19 PARP amino acid sequences from the Ensembl Main Database (Cunningham et al. 2022). PARP14rs2.1 was not chosen for analysis (NCBI 2022).

Since individual species analyses were unable to agree for some relationships we next used phylogenetic analysis to visualize relationships across and between species (Fig. 21). Analysis across all PARP species showed close relationship of DNA-PARPs (PARP1-3). PARP4, an undefined-PARP appears to be closely related to DNA-PARPs, although the confidence value is low (61). DNA-PARPs appear to be most-closely related to undefined-PARPs 4, 6, 8, and 16. PARP16, PARP8, and PARP6 appear to be even more closely related to each other. All Tankyrase-PARPs are closely related. Undefined-PARPs 10 may be related to Tankyrases based on these data. All Macro-PARPs are closely related to each other, but appear distant to the other PARPs. CCCH-PARPs along with undefined-PARP11 appears to group together. A tree for PARP fruit flies was not generated due to these species only having three members as phylogenetic analysis requires at least 4 species.

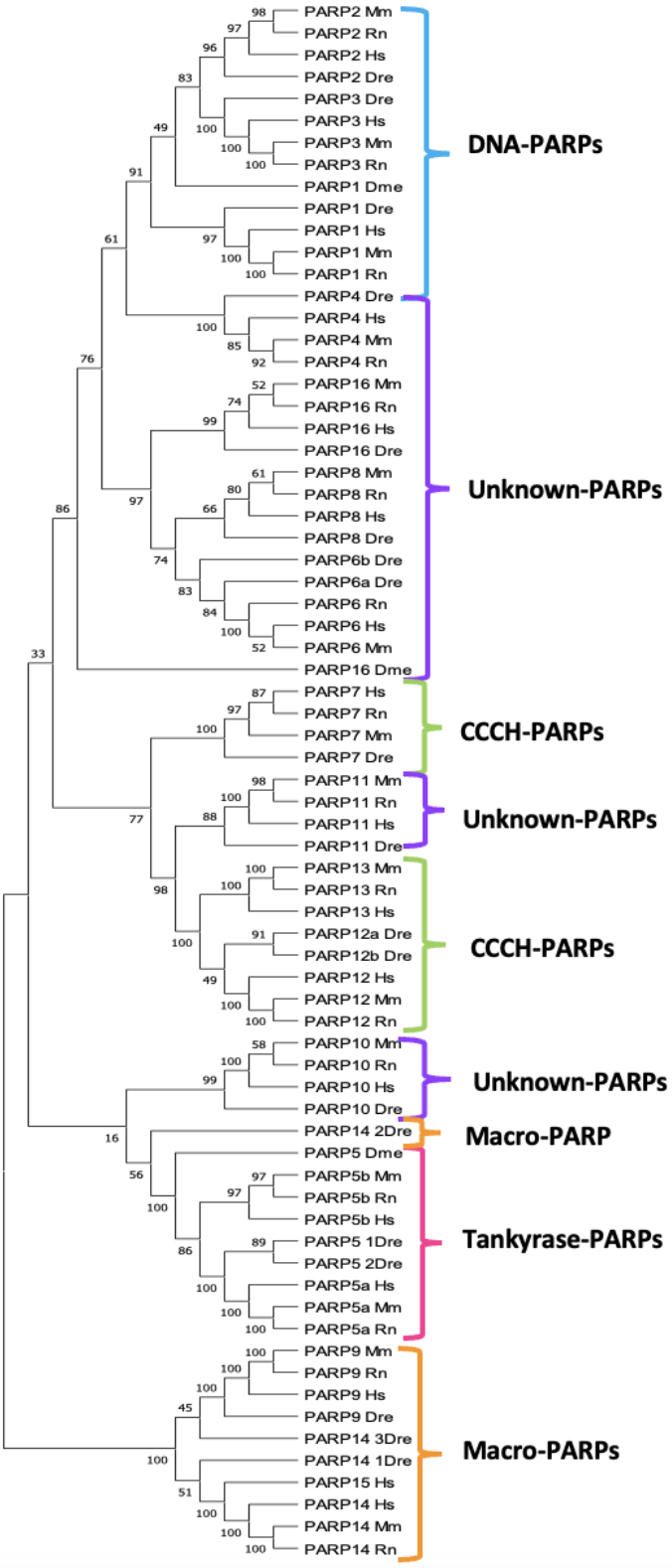

PARP2 Mm
PARP2 Rn
PARP2 Hs
PARP2 Dre
PARP3 Dre
PARP3 Hs
PARP3 Mm
PARP3 Rn
PARP1 Dme
PARP1 Dre
PARP1 Hs
PARP1 Mm
PARP1 Rn
PARP4 Dre
PARP4 Hs
PARP4 Mm
PARP4 Rn
PARP16 Mm
PARP16 Rn
PARP16 Hs
PARP16 Dre
PARP8 Mm
PARP8 Rn
PARP8 Hs
PARP8 Dre
PARP6b Dre
PARP6a Dre
PARP6 Rn
PARP6 Hs
PARP6 Mm
PARP16 Dme
PARP7 Hs
PARP7 Rn
PARP7 Mm
PARP7 Dre
PARP11 Mm
PARP11 Rn
PARP11 Hs
PARP11 Dre
PARP13 Mm
PARP13 Rn
PARP13 Hs
PARP12a Dre
PARP12b Dre
PARP12 Hs
PARP12 Mm
PARP12 Rn
PARP10 Mm
PARP10 Rn
PARP10 Hs
PARP10 Dre
PARP14 2Dre
PARP5 Dme
PARP5b Mm
PARP5b Rn
PARP5b Hs
PARP5 1Dre
PARP5 2Dre
PARP5a Hs
PARP5a Mm
PARP5a Rn
PARP9 Mm
PARP9 Rn
PARP9 Hs
PARP9 Dre
PARP14 3Dre
PARP14 1Dre
PARP15 Hs
PARP14 Hs
PARP14 Mm
PARP14 Rn
DNA-PARPs
Unknown-PARPs
CCCH-PARPs
Unknown-PARPs
CCCH-PARPs
Unknown-PARPs
Macro-PARP
Tankyrase-PARPs
Macro-PARPs

Figure. 21. Clustering by Amino Acid Similarity Shows *PARP* Genes Across Species Primarily Group by Function. The evolutionary history was determined using the Maximum-Likelihood method and JTT-matrix based model (Jones et al. 1992). The percentage of replicate trees in which the associated clade clustered together in the bootstrap test (100 replicates) are shown next to the branches (Felsenstein 1985). Initial tree(s) for the heuristic search were obtained by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. There were a total of 2958 positions in the final dataset and this analysis involved 71 amino acid sequences (Tables 3-6). PARP14rs2.1 member was not chosen for analysis.

Expression

Another way to investigate the evolutionary relationships among family members is to determine how well gene expression has been conserved among family members (Ocampo-Daza et al. 2011, Tostivint et al. 2014). This analysis examines regulatory conservation, such as functional aspects of promoter elements and enhancer activity, rather than sequence conservation in the coding region. In this first expression analysis, hierarchical clustering was used to examine pairwise co-expression of *PARP* genes (Obayashi et al. 2019). COXPRESdb, which includes data from microarray experiments of normal tissues and cell lines, was used to calculate and visualize co-expression of each gene pair as a heat map. Data are color scaled differently across species due to data retrieval times and number scales of data varied (Obayashi et al. 2019).

Human Macro-PARPs, *PARP9* and *PARP14* are the most co-expressed pair (Fig. 22). *PARP12*, a CCCH-PARP, was co-expressed with 9 and 14. *PARP10*, an undefined-PARP, and *PARP13* (*ZC3HAV1*) a CCCH-PARP, join the co-expression of *PARP9*, *PARP14*, and *PARP12* to a lesser degree. *PARP1* and *PARP2*, which are DNA-PARPs as well as *PARP5a* and *PARP5b*, the Tankyrase-PARPs shared some co-expression. Low or no co-expression is scattered across most of the other pairs of *PARP* genes. Similar to human, pairwise co-expression at high levels was observed between mouse Macro-PARPs, *PARP9* and *PARP14* (Fig. 23). *PARP9* and *PARP12* (a CCCH-PARP), as well as *PARP12* and *PARP14* displayed moderate co-expression levels. There is also moderate co-expression observed between *PARP9* and *PARP10* an undefined-PARP. *PARPs* 13/*ZC3HAV1* (a CCCH-PARP), *PARP11* (an undefined-PARP), *PARP4*, (another undefined-PARP), and *PARP3* (a

DNA-PARP) joined in co-expression to a lesser degree with those noted above. *PARP1* and *PARP2*, which are DNA-PARPs as well as *PARP5a* and *PARP5b*, the Tankyrase-PARPs shared some co-expression. The majority of the other pairs of *PARPs* had little or no co-expression.

Similar to human and mouse species in this study, pairwise co-expression at high levels can be observed between rat *PARP9* and *PARP14* (Fig. 24). In addition, pairs *PARP9* and *PARP12*, as well as *PARP12* and *PARP14* were moderately co-expressed. *PARP10* and *PARP13* join in co-expression with these noted above, just like in mouse. *PARP3* and *PARP4* also showed some co-expression with those noted above. Tankyrase-PARPs show some co-expression, as similarly seen in human and mouse. The majority of the other pairs of *PARPs* show little or no co-expression, similarly to mouse. Similar to human, mouse, and rat, zebrafish pairwise co-expression was highest between *PARP9* and *PARP14_1*/*PARP14rs1* (Fig. 25). *PARP10*, *PARP12a*, and *PARP14*_3/*PARP14rs4* showed some co-expression with each other, and the groups noted above. The majority of the other pairs of *PARPs* show little or no co-expression. In fruit flies *PARP* and *TNKS*, and *TNKS* and *PARP16* show minimal co-expression with one another (Fig. 26).

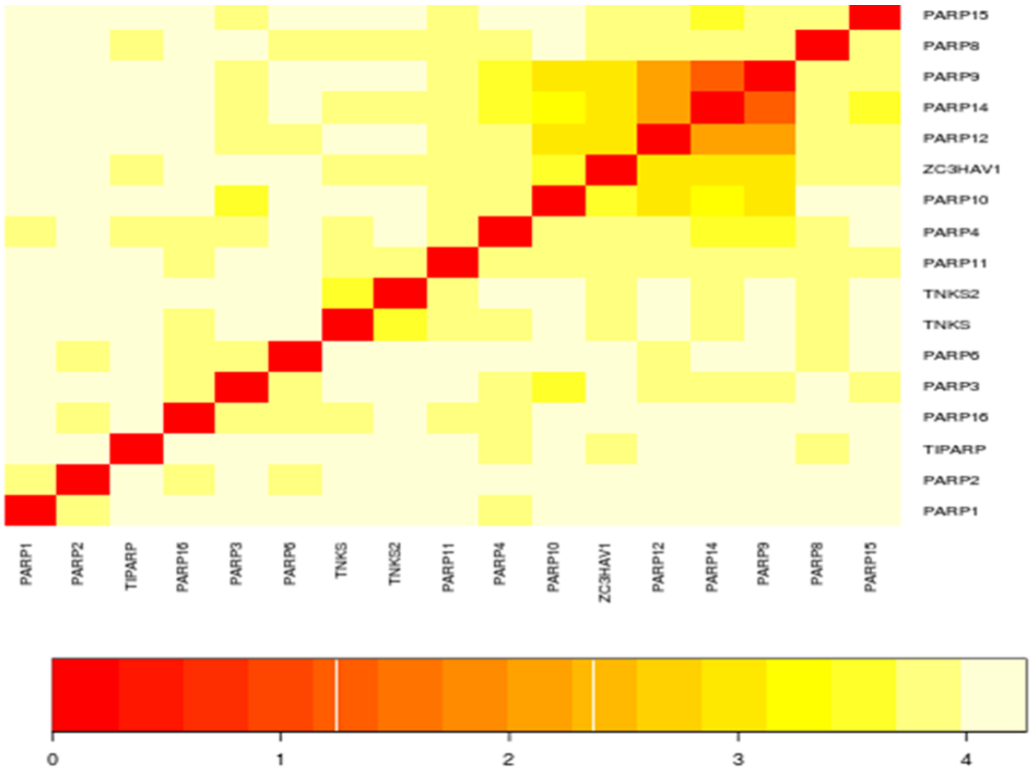


Figure. 22. *PARP9*, *PARP12*, and *PARP14* are Most Pairwise Co-Expressed Genes Among the Human *PARP* Gene Family. The *PARP* gene expression values from the NCBI microarray human data were clustered using Hcluster at the COXPRES database (Obayashi et al. 2019). Entrez gene ID numbers for these genes were obtained from NCBI. The linkage method for clustering was the average linkage method (Obayashi et al. 2019). Data are visualized in a heat map with high pairwise co-expression (0) being red to minimal pairwise co-expression (4) being light yellow. The following listed gene names are the equivalents for those not listed as their *PARP* gene names. *TIPARP*, is *PARP7*, *TNKS* and *TNKS2* are representative of *PARP5a* and *PARP5b*, and *ZC3HAV1* is *PARP13*.

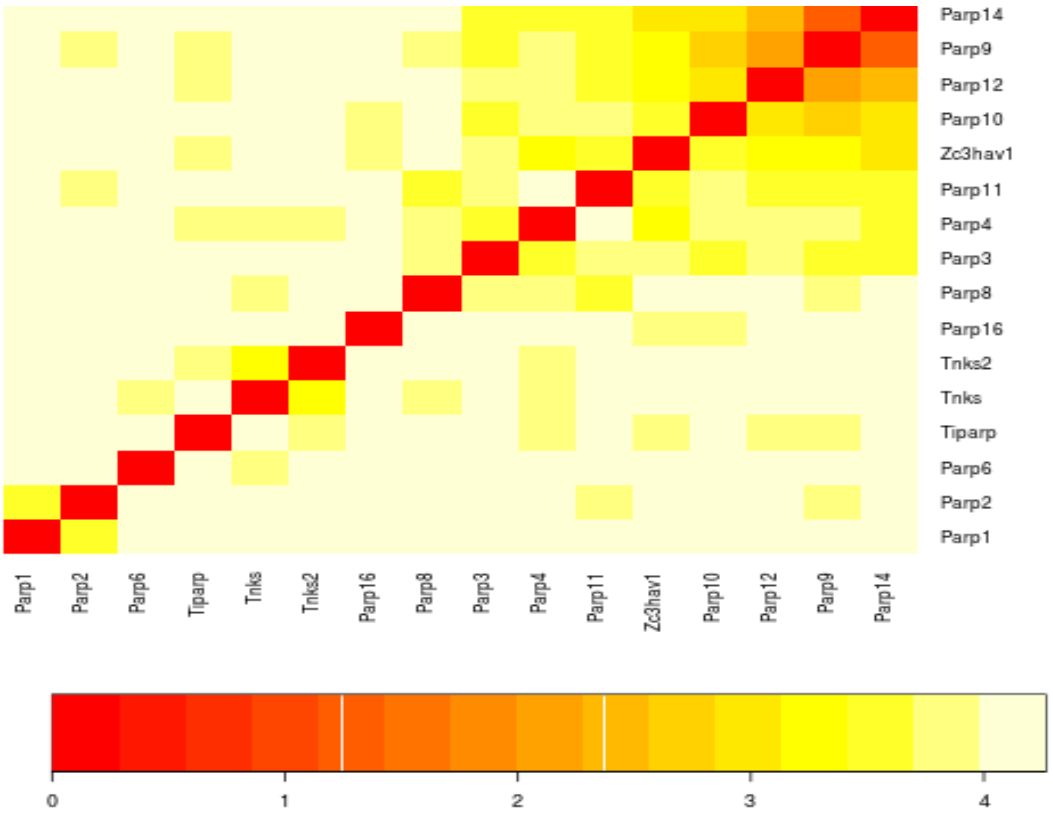


Figure. 23. *PARP9, PARP12, PARP14* are Most Pairwise Co-Expressed Genes Among the Mouse *PARP* Family Gene Members. The *PARP* gene expression values from NCBI microarray mouse data were clustered using Hcluster at the COXPRES database (Obayashi et al. 2019). The linkage method for clustering was the average linkage method (Obayashi et al. 2019). Data are visualized in a heat map with high pairwise co-expression (0) being red to minimal pairwise co-expression (4) being light yellow. The following listed gene names are the equivalents for those not listed as their *PARP* gene names. *Tiparp*, is *PARP7*, *Tnks* and *Tnks2* are representative of *PARP5a* and *PARP5b*, and *Zc3hav1* is *PARP13. PARP15* is not present in mouse.

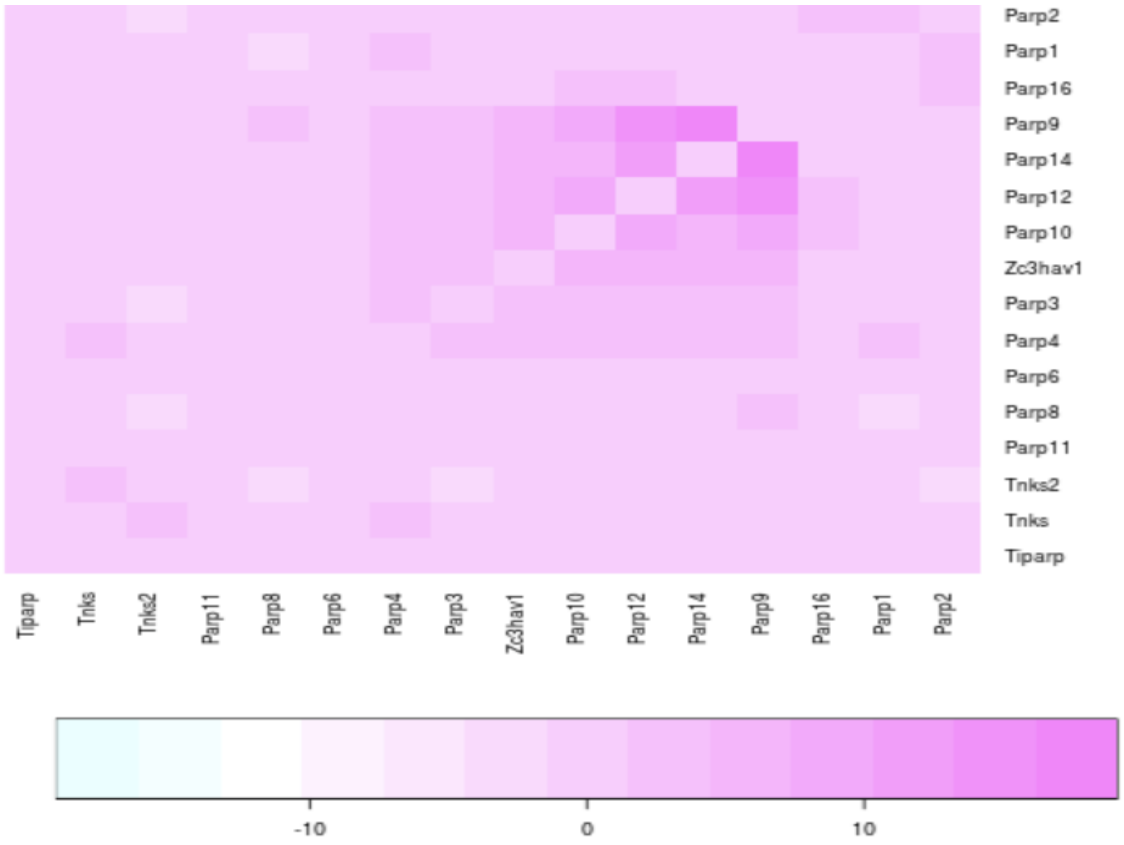


Figure. 24. *PARP9*, *PARP12*, and *PARP14* are Most Pairwise Co-Expressed Genes Among the Rat *PARP* Family Gene Members. The *PARP* gene expression values from NCBI microarray rat data were clustered using Hcluster at the database (Obayashi et al. 2019). The linkage method for clustering was the average linkage method (Obayashi et al. 2019). Data are visualized in a heat map with high pairwise co-expression (20) being dark pink to minimal pairwise co-expression (-0) light blue. The following listed gene names are the equivalents for those not listed as their *PARP* gene names. *TIPARP*, is *Parp7*, *TNKS* and *TNKS2* is representative of *Parp5a* and *Parp5b*, and *ZC3HAV1* is *Parp13*. *PARP15* is not present in rat.

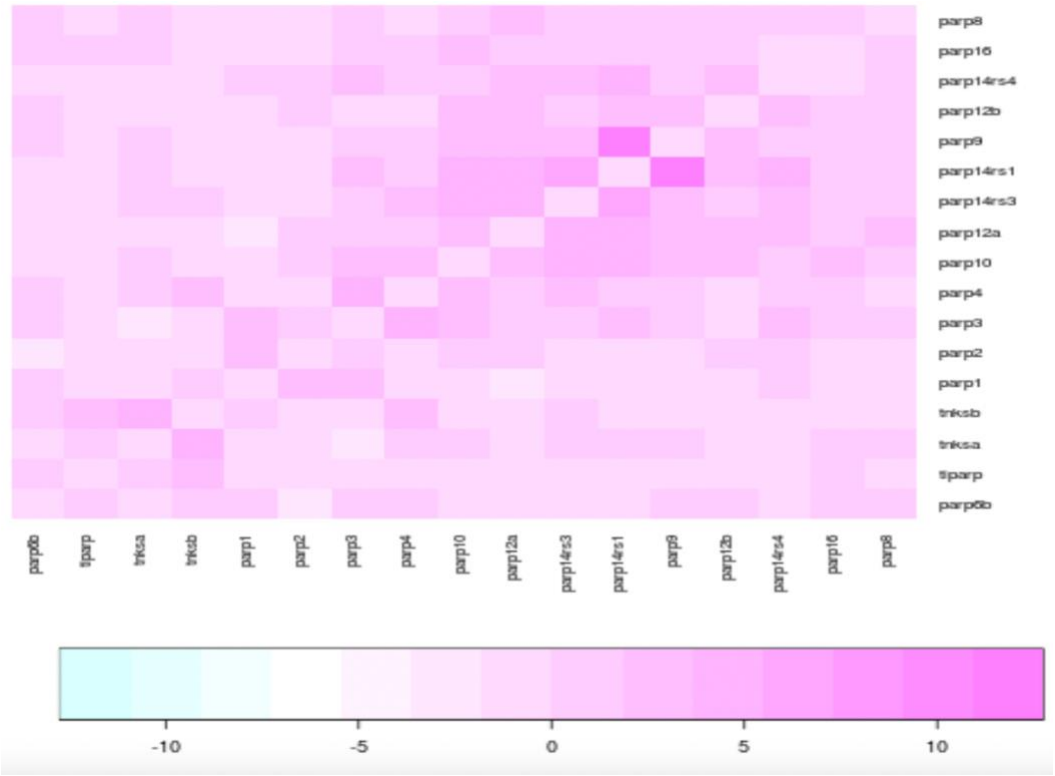


Figure. 25. *PARP9* and *PARP14*_1 are Most Pairwise Co-Expressed Genes Among the Zebrafish *PARP* Family Gene Members. The *PARP* gene expression values from NCBI microarray zebrafish data were clustered using Hcluster using the average linkage method (Obayashi et al. 2019). Data were visualized in a heat map with high pairwise co-expression (15) being dark pink to minimal pairwise co-expression (-0) light blue. Zebrafish *PARP* member 6a was not present in the database and *Parp11* has no Entrez Gene ID. *Tiparp* is *PARP7, Tnksa* (*PARP5_1*) and *Tnksb* (*PARP5_2*) are orthologs to *PARP5a*. *PARP13* and *PARP15* are not present in zebrafish. *PARP14rs2.1* member was not chosen for analysis.

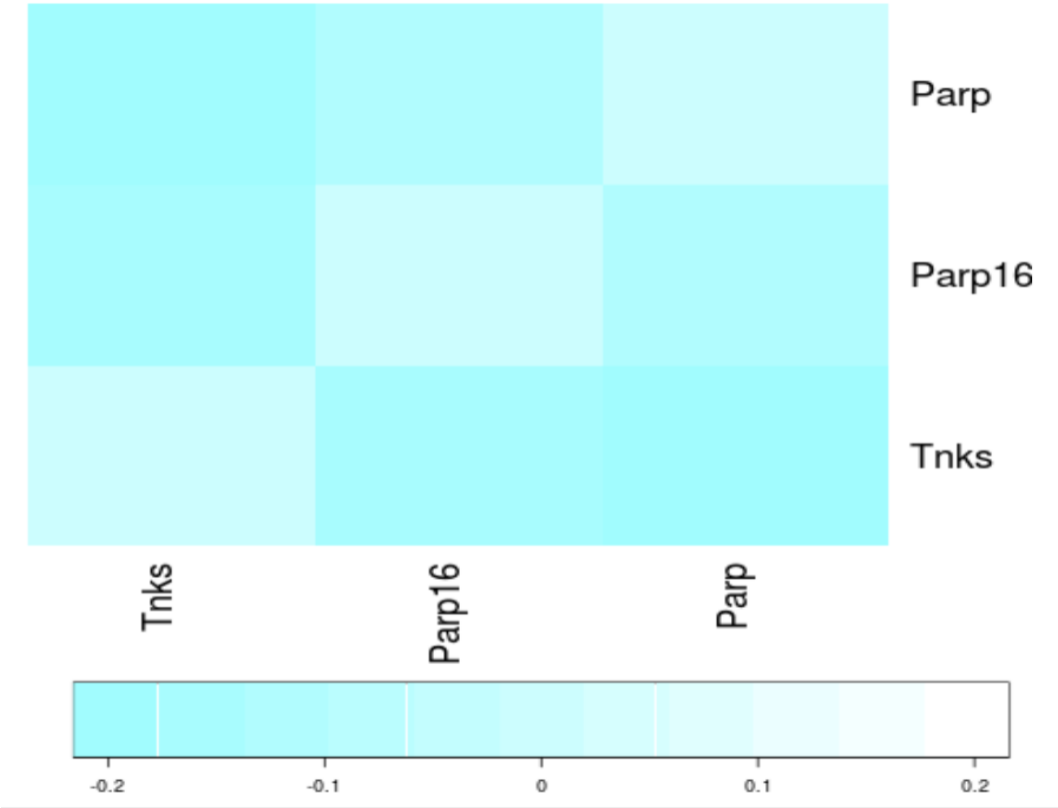
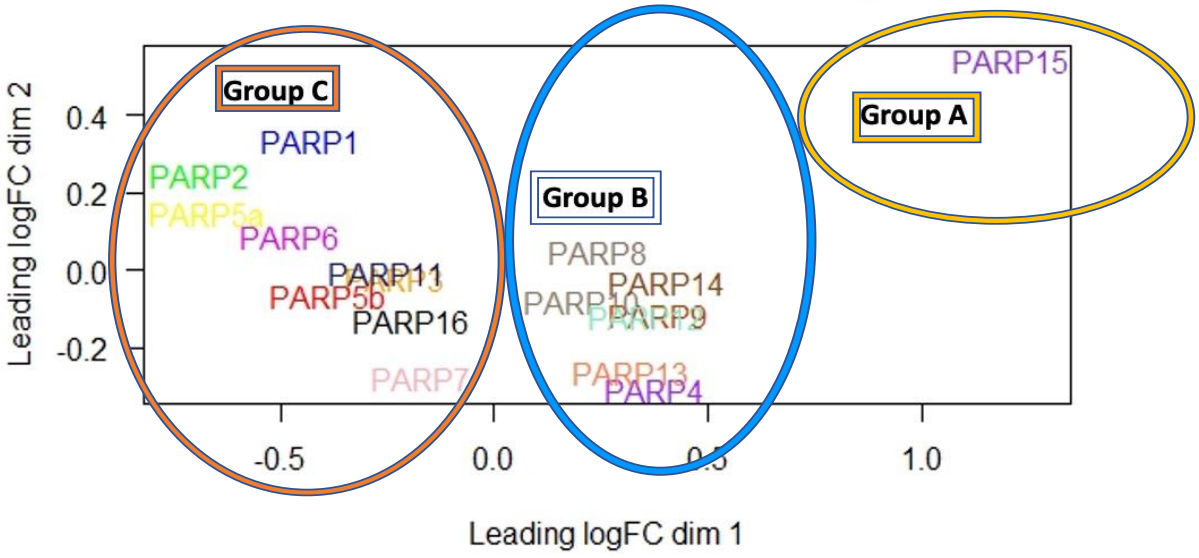


Figure. 26. *PARP*, *TNKS*, and *PARP16* show Minimal Pairwise Co-Expression Among the Fruit Fly *PARP* Family Gene Members. The *PARP* gene expression values from NCBI microarray fruit fly data were clustered using Hcluster at the COXPRES database (Obayashi et al. 2019). The linkage method for clustering was the average linkage method (Obayashi et al. 2019). Data are visualized in a heat map with high pairwise co-expression (0.2) being white to minimal pairwise co-expression (-0.2) light blue. The following listed gene names are the equivalents for those not listed as their *PARP* gene names. *Parp* is the equivalent to *PARP1* and *Tnks* is the equivalent to *PARP5a* and *PARP5b*. The other *PARP* gene family members are not present in fruit fly.

To gain a more thorough understanding of the conservation of gene expression for the human *PARP* gene family, as a set rather than just pairwise, a second approach was undertaken to cluster each gene's expression relative to all the other genes across normal tissues. This analysis utilized data from RNA-sequencing and contained only normal tissues (Genotype-Tissue Expression project) (GTEx 2021). Normalized gene expression values for each human *PARP* gene for 54 normal human tissues was used to calculate and produce a multi-dimensional scale plot and cluster dendrogram plot using R/R Studio (GTEx 2021, Loraine et al. 2015).

PARP genes in human tissues clustered into three groups: Group A, which include only *PARP15*, group B, which included *PARP13*, *PARP14*, *PARP9*, *PARP12*, *PARP10*, *PARP4*, and *PARP8*, and lastly group C, included *PARP11*, *PARP7*, *PARP16*, *PARP5b*, *PARP3*, *PARP6*, *PARP5a*, *PARP2*, and *PARP1* (Fig. 27 panel A). *PARP9* and *PARP12* and *PARP3* and *PARP11* are most similar across all tissues as they entirely overlap. Genes in this analysis appear to also have clustered somewhat according to grouped functions (Fig. 27 panel B). In group A, *PARP15* is a Macro-PARP, while group B contains CCCH-PARPs, undefined-PARPs, and Macro-PARPs. Group C has the remaining undefined-PARPs, DNA-PARPs, Tankyrase-PARPs, and a CCCH-PARP member.

A


Group C
Group B
Group A
PARP1
PARP2
PARP5a
PARP6
PARP11
PARP5b
PARP16
PARP7
PARP8
PARP14
PARP10
PARP9
PARP13
PARP4
PARP15
Leading logFC dim 2
Leading logFC dim 1
0.4
0.2
0.0
-0.2
-0.5
0.0
1.0

B

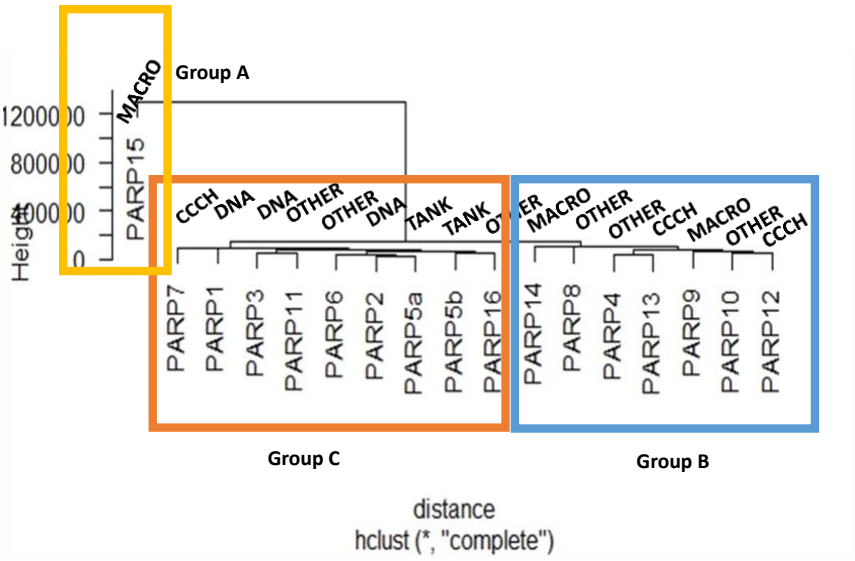
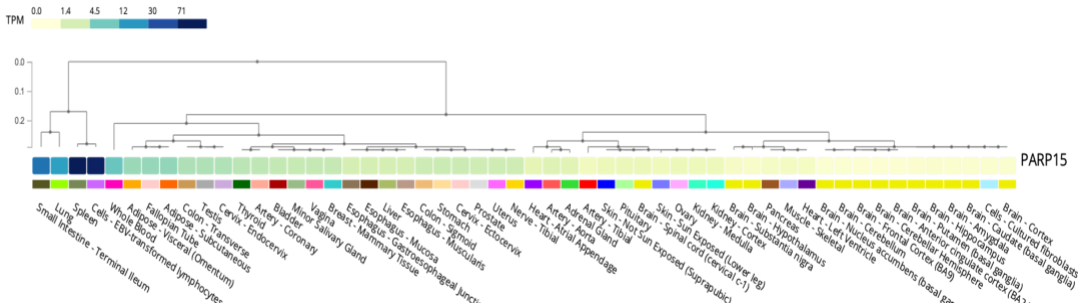

Group A
MACRO
PARP15
CCCH
DNA
DNA
OTHER
OTHER
DNA
TANK
TANK
MACRO
OTHER
OTHER
CCCH
MACRO
OTHER
CCCH
PARP7
PARP1
PARP3
PARP11
PARP6
PARP2
PARP5a
PARP5b
PARP16
PARP14
PARP8
PARP4
PARP13
PARP9
PARP10
PARP12
Group C
Group B
distance
hclust (*, "complete")

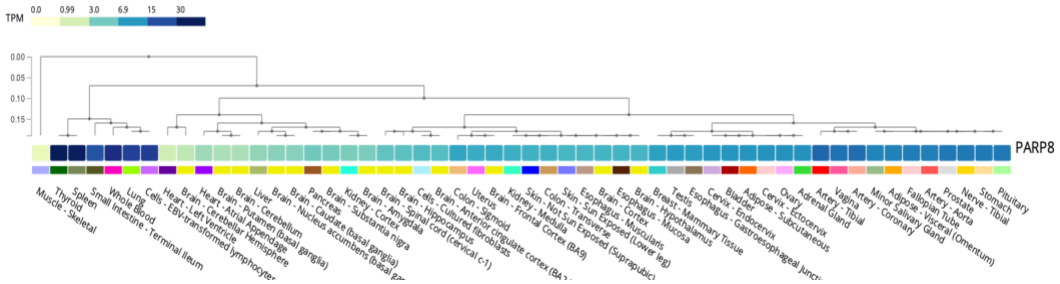
Figure. 27. MDS and Cluster Dendrogram Analysis of Human *PARP* Gene Expression Across 54 Normal Tissues Shows Three Similarly Expressed *PARP* Groups. In panel A, the median gene expression levels for each *PARP* gene was obtained from the Genotype-Tissue Expression (GTEx) project for 54 normal tissues collected across 948 donors (GTEx 2021). Clustering of gene expression was calculated then visualized using a Multi-Dimensional Scale (MDS) plot using R/R studio (Loraine et al. 2015). Each of the *PARP* genes are presented in three large, clustered groups (group A: colored yellow, group B: colored blue, and group C: colored orange) according to similar expression and are grouped as follows: *PARP15* forms group A. *PARP8*, *PARP14*, *PARP10*, *PARP9*, *PARP12*, *PARP13*, and *PARP4* form group B. *PARP1*, *PARP2*, *PARP5a*, *PARP6*, *PARP11*, *PARP3*, *PARP5b*, *PARP16*, and *PARP7* form group C. Similarly in panel B, Clustering of gene expression was calculated then visualized using a Cluster Dendrogram plot using R/R studio (Loraine et al. 2015). Each of the *PARP* genes are presented in three large, clustered groups (groups A, B, and C) according to similar expression and are grouped as follows: *PARP15* (Macro-PARP) forms group A. *PARP4*, *PARP8*, and *PARP10* (Other-PARPs), *PARP9*, *PARP14* (Macro-PARPs), and *PARP12* and *PARP13* (CCCH-PARPs) form group B. *PARP1*, *PARP2*, and *PARP3* (DNA-PARPs), *PARP5a*, *PARP5b*, (Tankyrase-PARPs), *PARP6*, *PARP11*, *PARP16*, (Other-PARPs), and *PARP7* (CCCH-PARP) form group C.

To gain an idea of how these three groups differed in expression, across the normal tissues, expression from one gene from each group was visualized using a heat map using GTEx (GTEx 2022). Group A has one member, which is *PARP15*. The highest expression levels of *PARP15* were observed in both the spleen and EBV-transformed lymphocytes (Fig. 28 panel A). The lowest expression levels of *PARP15* were in the brain. *PARP8*, a representative gene from group B, had the highest expression in the thyroid and spleen (Fig. 28 panel B). The lowest expression for *PARP8* was in the skeletal muscle. Overall, *PARP3*, a representative gene from group C was expressed in each of the tissues (Fig. 28 panel C). The highest expression was in the adrenal gland and thyroid. High expression levels of *PARP3* were also in tissues such as the lung, spleen, liver, and uterus. The lower expression levels of *PARP3* were brain tissue groups. The pancreas and whole blood tissues had the lowest *PARP3* expression.

A



B



C

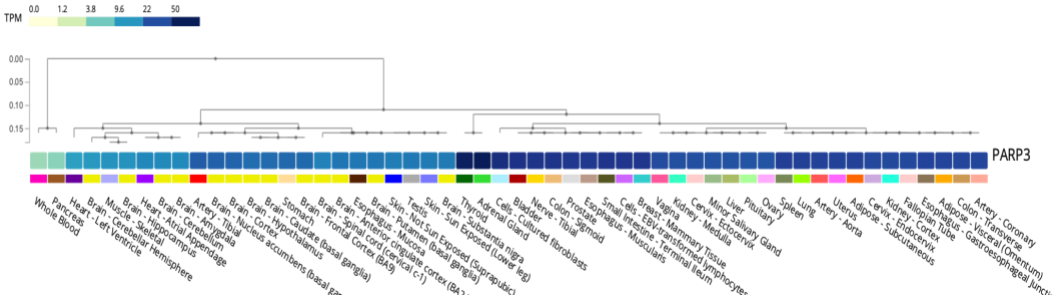


Figure 28. Expression of Example Genes. Gene expression was visualized at GTEx (2022) as a heat map quantified as transcript per million (TPM).Group A representative: *PARP15.* The TPM score for this gene is on a scale of 0.0 (light yellow) to 71 (dark blue). Group B representative: *PARP8*. The TPM score for this gene is on a scale of 0.0 (light yellow) to 30 (dark blue). Group C representative: *PARP3*. The TPM score for this gene is on a scale of 0.0 (light yellow) to 50 (dark blue).

Gene Order and Synteny

Another way to evaluate evolutionary changes for a gene family is to examine gene order and synteny changes among species to identify chromosomal conservation, divergence, deletions, insertions, and rearrangements (Maynard et al. 2014). Each *PARP* gene was identified, surrounding genes were noted, and included the direction of transcription (NCBI, Ensembl Main, Metazoan Ensembl, Tables 7-11, JAX Synteny Browser Kolishovski et al. 2019). Each region was mapped with the human *PARP* gene as the anchoring gene (NCBI, Ensembl Main, Metazoan Ensembl, Tables 7-11, JAX Synteny Browser Kolishovski et al. 2019).

PARP1 is located on chromosome 1 in both mice and human. Order and direction of transcription for surrounding genes *ITPKB*, *STUM*, *LIN9*, and *MIXL1* was maintained for mouse, rat, and human (Fig. 29, NCBI, Tables 7-11). Only *LIN9* and *MIXL1* are present in zebrafish near its *PARP1* gene. Genes found in this region that were not present in other species were located in their individual genomes to ensure these genes were not simply deleted. The *ITPKB* gene in human is noted as the ortholog to the *IP3K2* gene in *Drosophila* which is located on chromosome X. The *STUM* gene is located on chromosome 2R, while the *LIN9* in human is an ortholog to the *ALY* gene located on chromosome 3L. *MIP130* is also an ortholog to the human *LIN9* gene and is located on chromosome X. *MIXL1* gene is not present in *Drosophila*.

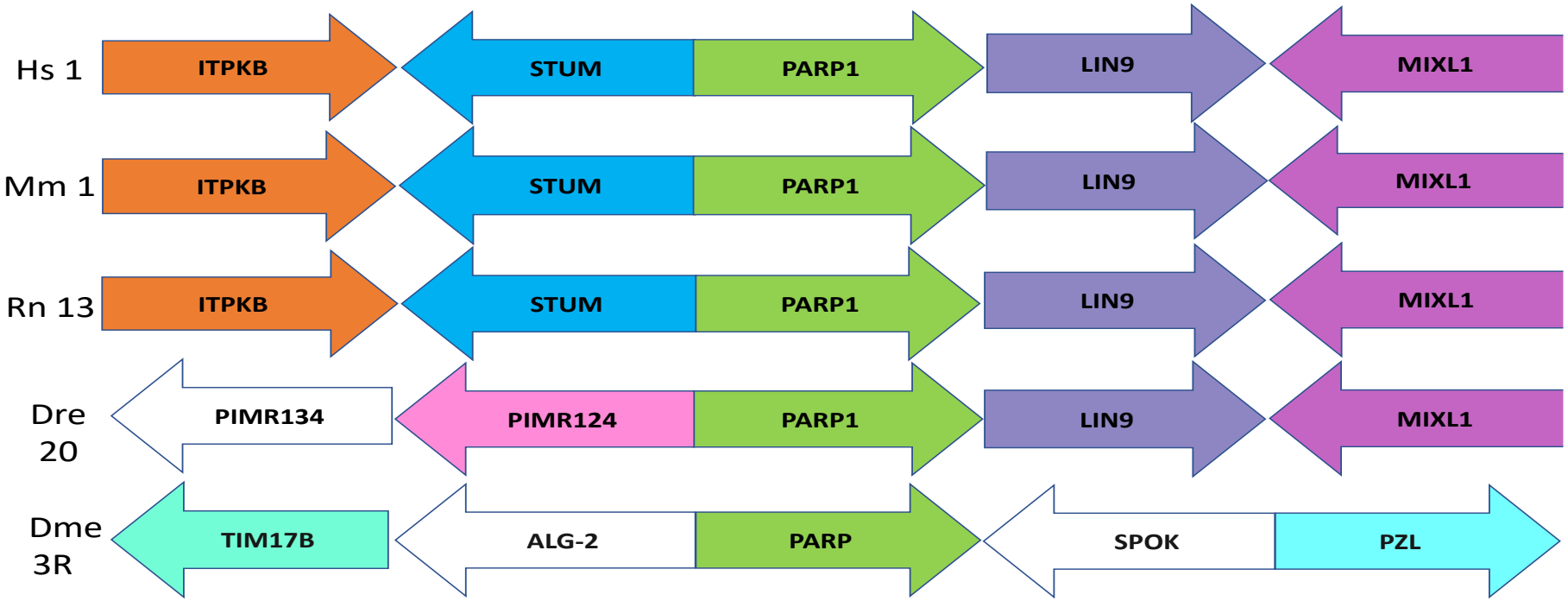


Figure. 29. *PARP1* Chromosomal Context Shows Similar Conserved Gene Order for Four Out of Five Species in this Study. The NCBI Genome Database was used to identify *PARP1's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts.

In similar fashion, *PARP2* was located on chromosome 14 in both human and mouse. The same gene order in mouse and rat has been maintained when compared to humans including surrounding genes *TTC5*, *CCNB1IP1*, *TEP1*, and *KLHL33* (Fig. 30, NCBI, Tables 7-11). In zebrafish the order was no longer maintained, as the other genes (*SDR39U1*, *METTL17*, *SI:CH211-284O19.8* (ortholog to the human *GRINA* gene), *ZGC:66427*) surround *PARP2*. *METTL17* in human and mouse was located further down on chromosome 14 and in rat it was located further down on chromosome 15. In zebrafish the *TTC5* and *CCNB1LIP1* genes were located on chromosome 7. The *SI:CH211-63P21.8* gene is an ortholog to human gene *KLHL33* and is also located on chromosome 7.

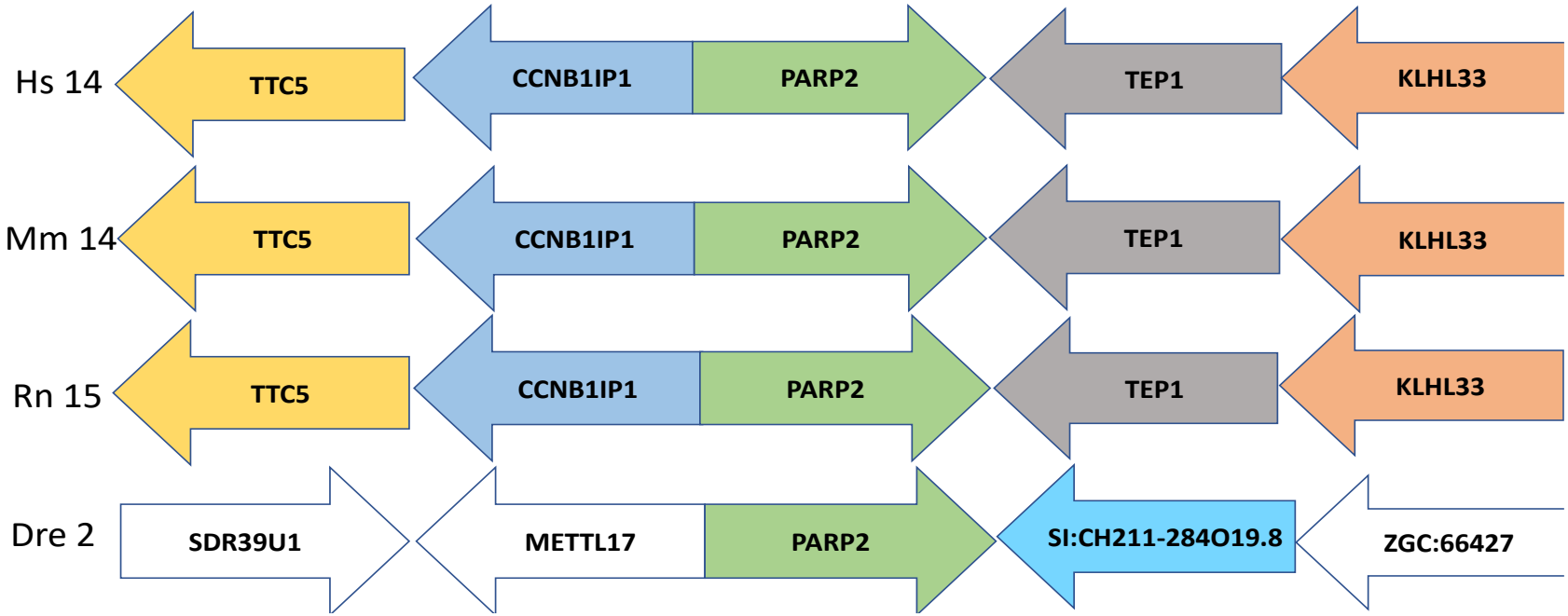


Figure. 30. *PARP2* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study.

The NCBI Genome Database was used to identify *PARP2's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP2* gene is absent in *Drosophila*.

PARP3 is found on chromosome 3 in human, 9 on mouse, 8 on rat, and 6 on zebrafish. The gene order *IQCF1*, *RRP9*, *GPR62*, and *PCBP4* was the same across human, mouse, and rat but differed in zebrafish (Fig. 31, NCBI, Tables 7-11). *TEX264A*, *GRM2A*, *GPR61L*, and *CAV3* are the nearest genes in this species. In zebrafish *RRP9* is located on chromosome 11, *IQCF1* was not present, and the *SI:CH211-213O11.11* gene is an ortholog to the human gene *GPR62*, which was also located on chromosome 11. The *PCBP4* gene is located on chromosome 22.

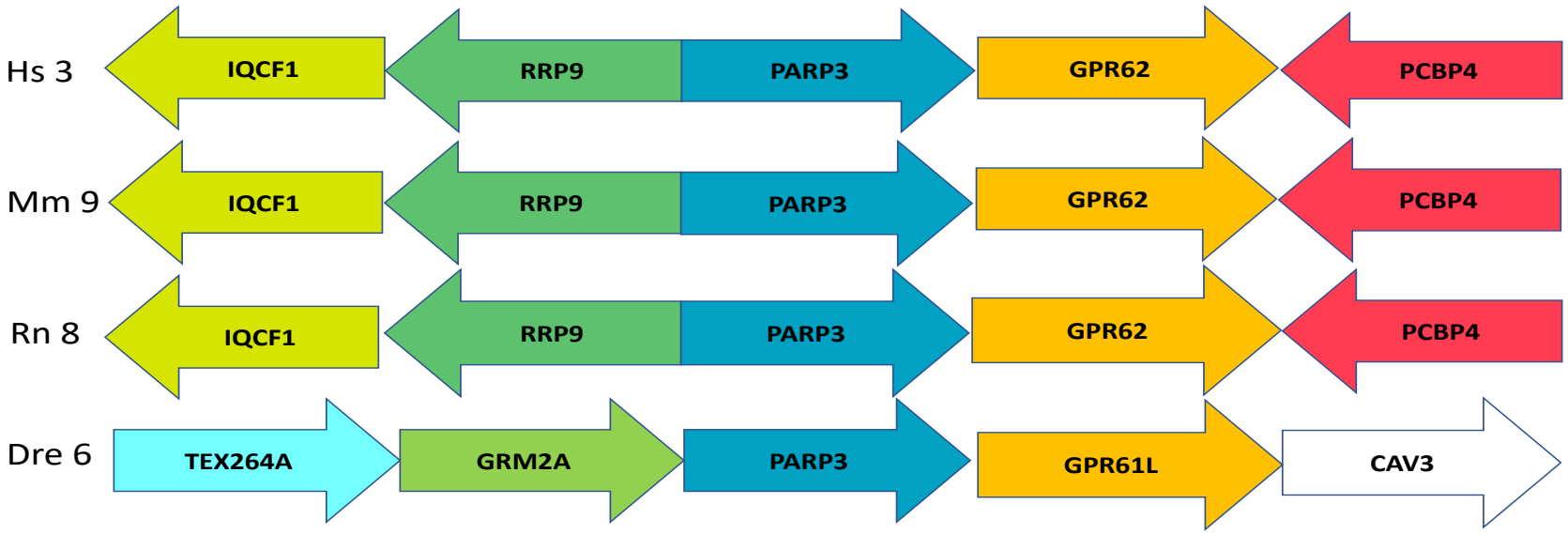


Figure. 31. *PARP3* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study. The NCBI Genome Database was used to identify *PARP3's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP3* gene is absent in *Drosophila*.

For *PARP4*, the surrounding genes *RNF17*, *ATP12A*, *CENPJ*, *MPHOSPH8*, and *PSPC1* were in identical order in mouse and rat but differed when compared to human. *RNF17* and *ATP12A* are inverted on chromosome 13, indicating a chromosomal inversion event (Fig. 32, NCBI, Tables 7-11). *CENPJ* was located further down chromosome 13 in human and *C1QTNF9* and *SPATA13* were the other genes that surround *PARP4*. *MPHOSPH8* and *PSPC1* were located further down chromosome 13 in human. In mouse *C1QTNF9* was located on chromosome 14, *SPATA13* was located on chromosome 14, and *SPATA13* was located in front of *C1QTNF9*. In rat *C1QTNF9* and *SPATA13* was located on chromosome 15, and similarly *SPATA13* was located in front of *C1QTNF9*. In zebrafish there were different surrounding genes compared to the other species (*LOC103911649*, *IGSF3*, *TRIM45*, and *VTCN1*). In this species the *ATP12A* gene was not present, *RNF17* was located on chromosome 9, and *CENPJ* was also located on chromosome 9. *C1QTNF9* was located on chromosome 24 and *SPATA*13 was located on chromosome 24.

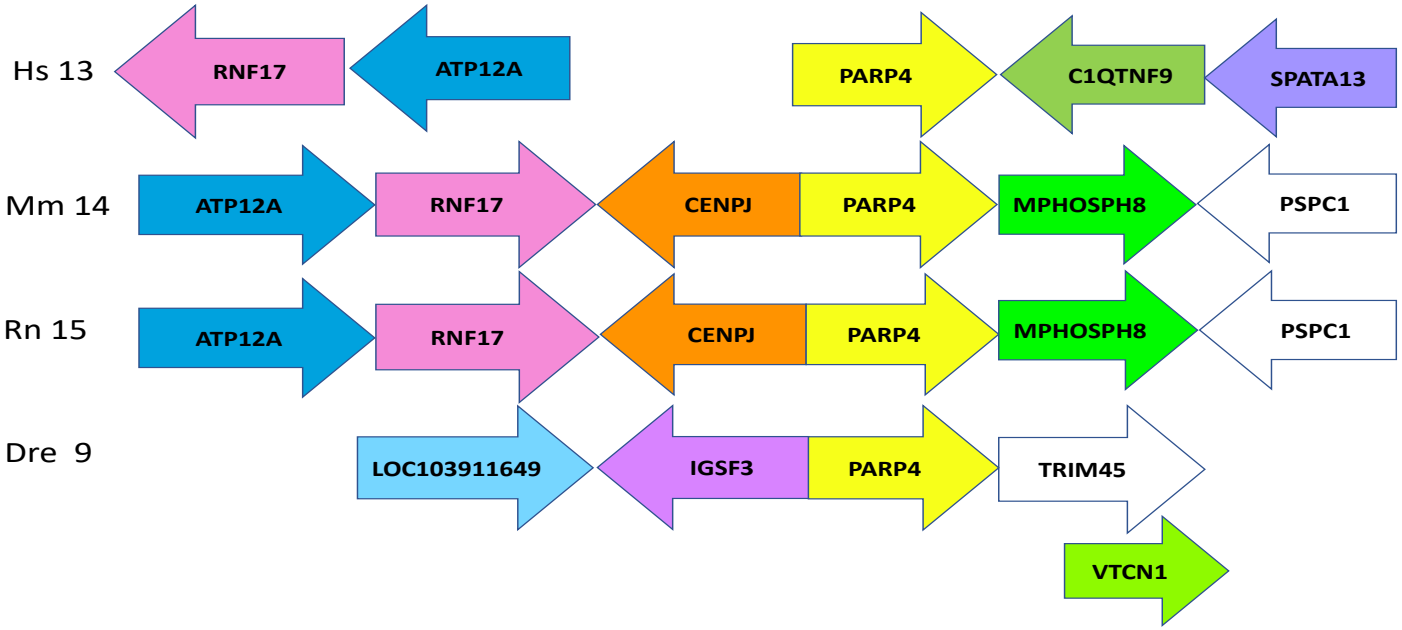
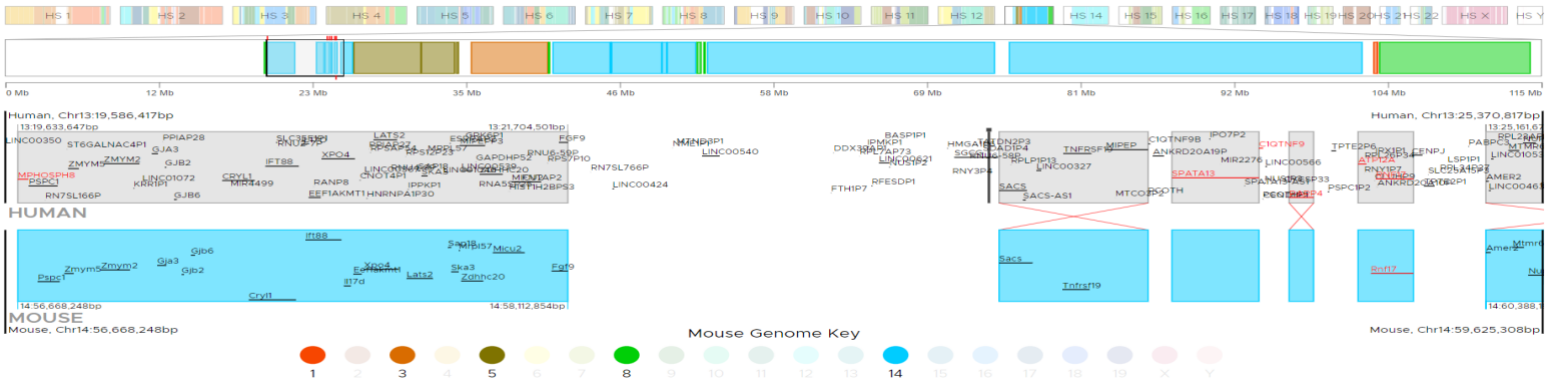

Hs 13
RNF17
ATP12A
PARP4
C1QTNF9
SPATA13
Mm 14
ATP12A
RNF17
CENPJ
PARP4
MPHOSPH8
PSPC1
Rn 15
ATP12A
RNF17
CENPJ
PARP4
MPHOSPH8
PSPC1
Dre 9
LOC103911649
IGSF3
PARP4
TRIM45
VTCN1

Figure. 32. *PARP4* Chromosomal Context Shows Similar Conserved Gene Order for Three Out of Five Species in this Study.

The NCBI Genome Database was used to identify *PARP4's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP4* gene is absent in *Drosophila*.

Synteny analysis using the JAX Synteny browser allowed a broader look at this region. *RNF17* and *ATP12A* showed a chromosomal inversion, as well as *PARP4* itself (Fig. 33, JAX Synteny Browser Kolishovski et al. 2019). When mouse and human were compared highlighted by blocks of synteny, the genomic block containing *PARP4* was found inverted relative to its neighboring synteny blocks (Fig. 33 A). The genomic block containing human *C1QTNF9* and *SPATA13* was located on a more distant block of mouse chromosome 14 (Fig. 33 A, B). Human *MPHOSPH8* and *PSPC1* were more distantly located from *PARP4* on chromosome 13, but located directly next to *PARP4* on mouse chromosome 14 (Fig. 33 A, B).

A



B

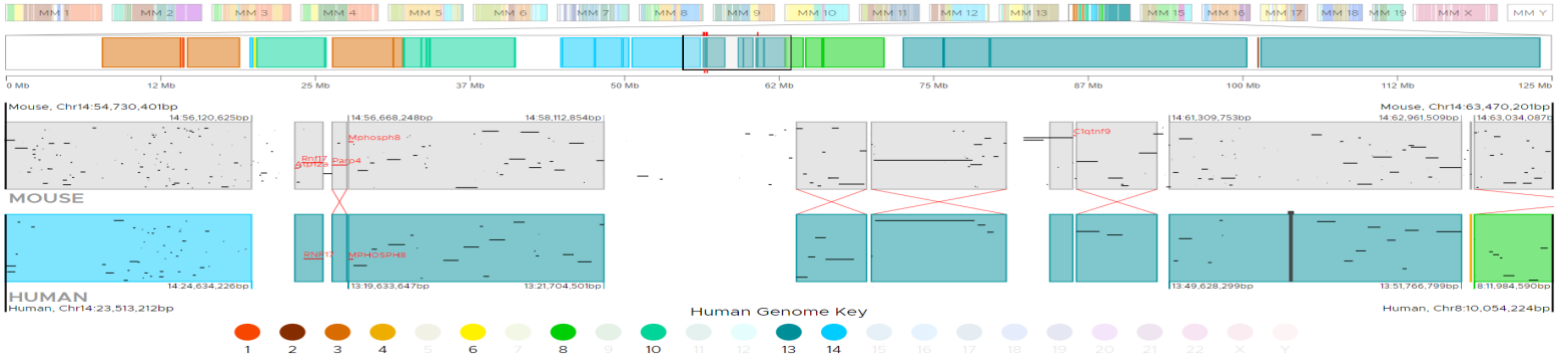


Figure. 33. Human-Mouse Synteny surrounding *PARP4*. Gene order and changes visualized in the JAX Synteny Browser (Kolishovski et al. 2019) with A) Human compared to Mouse and B) Mouse compared to Human. Genes highlighted in red are *RNF17*, *MPHOSPH8*, *PARP4*, *ATP12A*, *C1QTNF9*, and *SPATA13*.

In mouse and rat the chromosome context surrounding *PARP5a* relative to humans is not identical (Fig. 34, NCBI and Ensembl, Tables 7-11). *ERI1* and *PPP1R3B* genes in human and mouse have maintained order on chromosome 8. Differences include the genes *MSRA* and *PRSS51* in human but *DUSP4* and *SARAF* in mouse and rat. Human *DUSP4* and *SARAF* genes were located further upstream. In mouse and rat *MSRA* and *PRSS51* are located on chromosomes 14 and 15 respectively. Synteny analysis showed that *PARP5a*/*TNKS*, *MSRA*, and *PPP1R3B* in human and mouse are found in syntenic genomic blocks (chromosome 8 for both) but the next genomic block in human, which carries *MSRA*, is syntenic with mouse chromosome 14 (Fig. 35 A, JAX Synteny Browser Kolishovski et al. 2019). Human *DUSP4* and *SARAF* are found downstream of *PARP5a* in a block that is also syntenic with mouse chromosome 8 (Fig. 35 B).

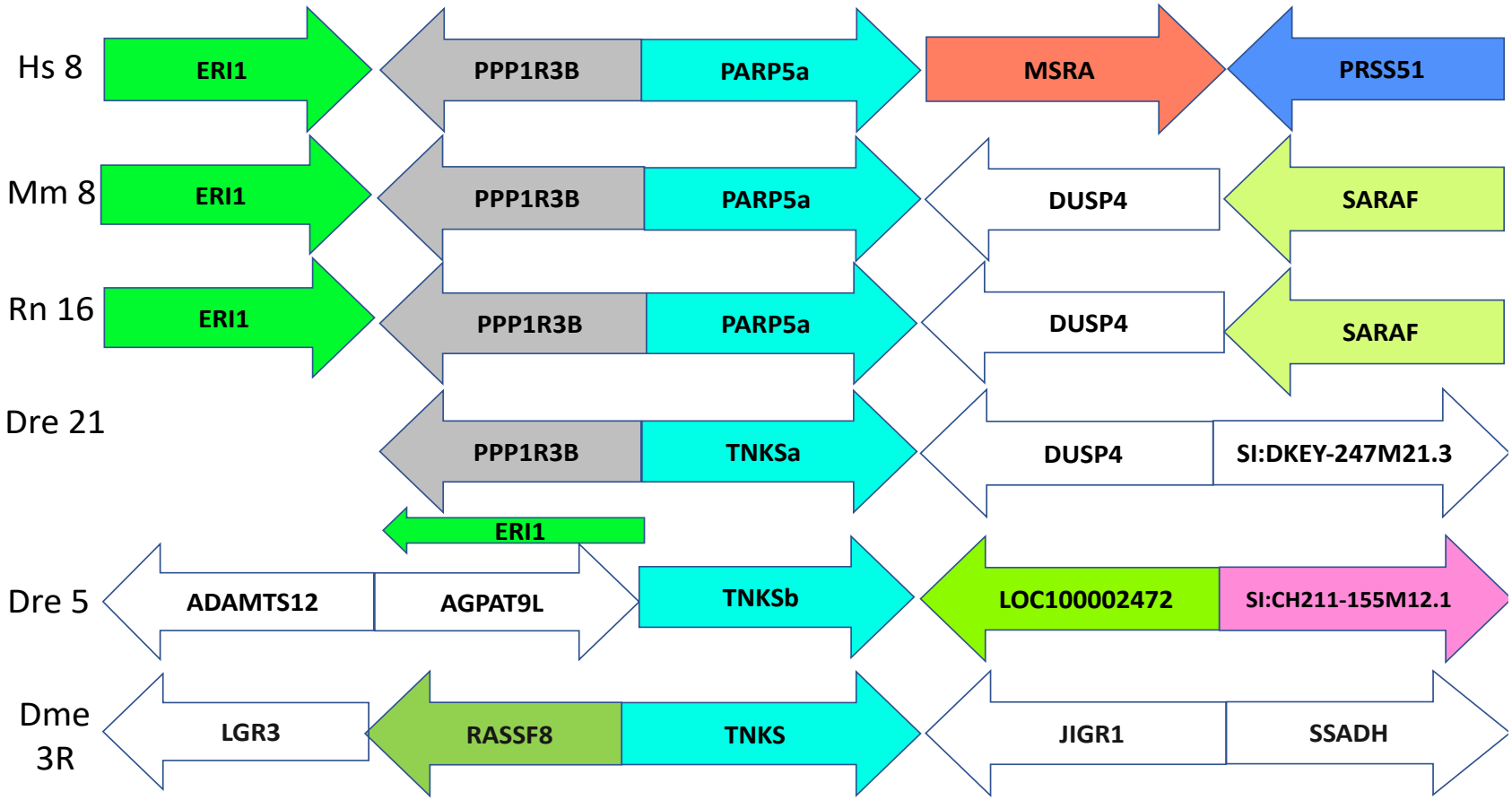
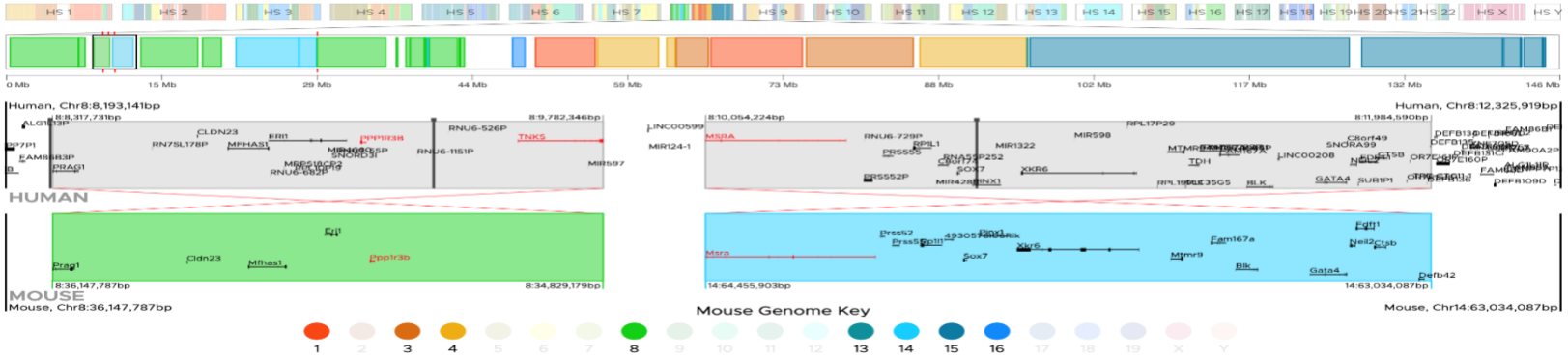

Hs 8
ERI1
PPP1R3B
PARP5a
MSRA
PRSS51
Mm 8
ERI1
PPP1R3B
PARP5a
DUSP4
SARAF
Rn 16
ERI1
PPP1R3B
PARP5a
DUSP4
SARAF
Dre 21
PPP1R3B
TNKSa
DUSP4
SI:DKEY-247M21.3
ERI1
Dre 5
ADAMTS12
AGPAT9L
TNKSb
LOC100002472
SI:CH211-155M12.1
Dme 3R
LGR3
RASSF8
TNKS
JIGR1
SSADH

Figure. 34. *PARP5a* Chromosomal Context Shows Similar Conserved Gene Order for Four Out of Five Species in this Study. The NCBI Genome and Ensembl Databases were used to identify *PARP5a's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (Cunningham et al. 2022, NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. In zebrafish both *TNKSa* and *TNKSb* are listed as orthologs to the human *PARP5a* gene (Cunningham et al. 2022). The *TNKS* gene in *Drosophila* (Dme), is listed as an ortholog to both human *PARP5a* and *PARP5b* as it is unclear which human gene it is more closely related to (Cunningham et al. 2022, Yates et al. 2022).

A



B

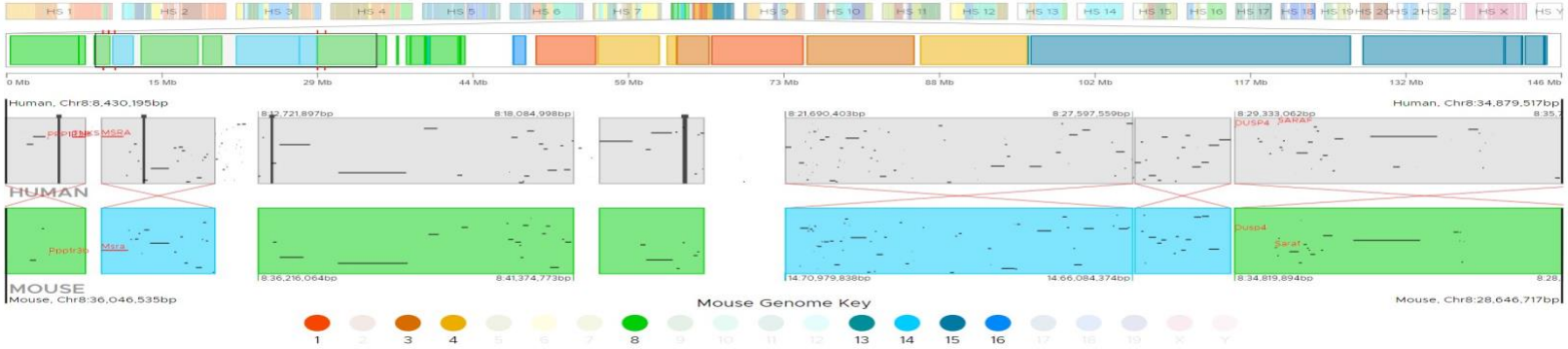


Figure. 35. Human-Mouse Synteny surrounding *PARP5a*. Gene order and changes visualized in the JAX Synteny Browser (Kolishovski et al. 2019) with A) and B) Human compared to Mouse. Genes highlighted in red are *DUSP4*, *SARAF*, *PPP1R3B*, *PARP5a/TNKS*, and *MSRA*.

In zebrafish there are two human *PARP5a* ortholog members, *TNKSa* and *TNKSb*. On chromosome 21 *ERI1* is overlapping with *PPP1R3B,* and *DUSP4* is located on the other side of *TNKSa*. In *TNKSa* in zebrafish *MSRA* is located on chromosome 20 and *PRSS51* is not present. *TNKSb* has diverged from the other *PARP5* members overtime, as the genes *ADAMTS12*, *AGPAT9L*, *LOC100002472*, and *SI:CH211-155M12.1* surrounded this gene. This is also true for fruit fly, as the genes *LGR3*, *RASSF8*, *JIGR1*, and *SSADH* on chromosome 3R surround *TNKS* in this species. In *Drosophila ERI1* is not present, *PPP1R3B* in human is the ortholog to the *Gbs-70E* gene in this species and is located on chromosome 3L. *MSRA* is located on chromosome 3L and *PRSS51* is not present. *HECTD2, PPP1R3C, FGFBP3, BTAF1* are the genes that surround *PARP5b* and have maintained order overtime relative to human, as mouse and rat have the identical context (Fig. 36, NCBI and Ensembl, Tables 7-11). Relative to the *TNKS* gene in fruit fly, it is an ortholog to both the human *PARP5a* and *PARP5b* genes and is why it is listed on both of their gene contexts (Fig. 34, Fig. 36).

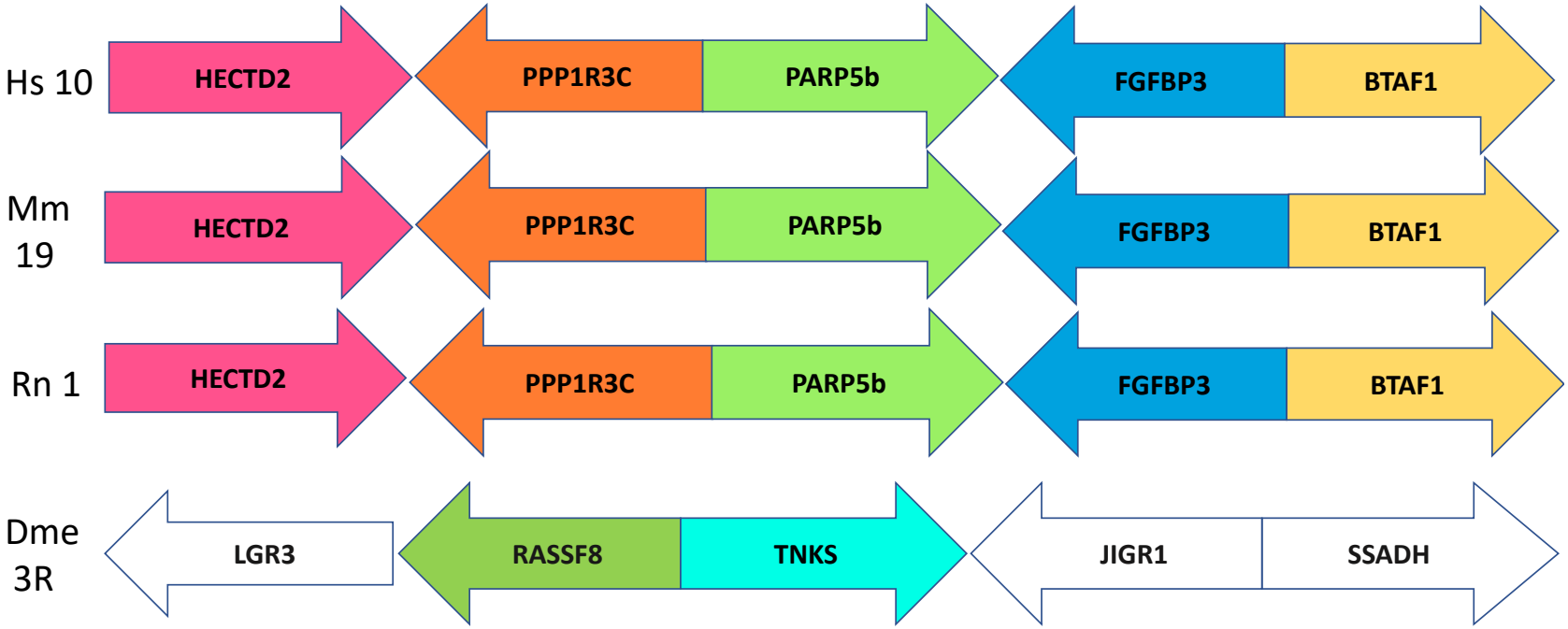

Hs 10
HECTD2
PPP1R3C
PARP5b
FGFBP3
BTAF1
Mm 19
HECTD2
PPP1R3C
PARP5b
FGFBP3
BTAF1
Rn 1
HECTD2
PPP1R3C
PARP5b
FGFBP3
BTAF1
Dme 3R
LGR3
RASSF8
TNKS
JIGR1
SSADH

Figure. 36. *PARP5b* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study.

The NCBI Genome Database was used to identify *PARP5b's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order context (NCBI 2022). The *TNKS* gene in *Drosophila* (Dme), is listed as an ortholog to both human *PARP5a* (*TNKS*) and *PARP5b* (*TNKS2*) genes (Cunningham et al. 2022). In zebrafish both *TNKSa* and *TNKSb* genes are listed as orthologs to the human *PARP5a*, which is why they are not listed in this gene's chromosome context.

Relative to the human *PARP6* gene both mouse and rats have identical context, which include *HEXA*, *CELF6*, *PKM*, and *GRAMD2-GRAMD2A* (Fig. 37, NCBI and Ensembl, Tables 7-11). The genes that surround *PARP6a* in zebrafish on chromosome 18 include *DHDH.2*, *PKMA*, *SI:DKEY-205H23.1*, and *FAM219B*. The genes surrounding *PARP6b*, are located on chromosome 25, and include *HCN4L*, *PKMB*, *RPP25A*, and *COX5AA*. The *HEXA* gene in zebrafish is an ortholog to the human *HEXA* gene and is located on chromosome 25. *CELF6* is located on chromosome 18, and the *GRAMD2AA* gene, which is an ortholog to the human *GRAMD2A* gene is located on chromosome 7.

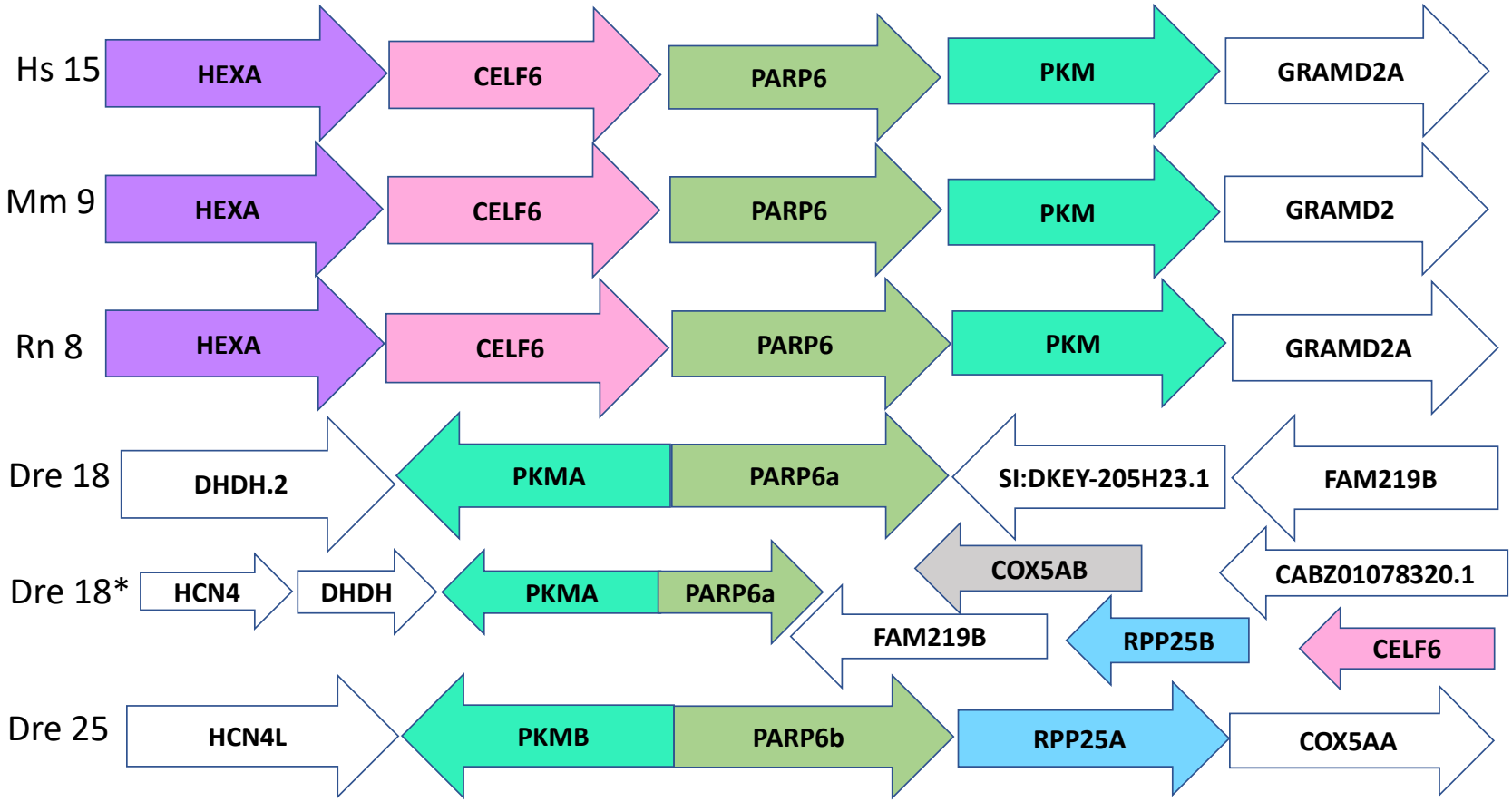

Hs 15
HEXA
CELF6
PARP6
PKM
GRAMD2A
Mm 9
HEXA
CELF6
PARP6
PKM
GRAMD2
Rn 8
HEXA
CELF6
PARP6
PKM
GRAMD2A
Dre 18
DHDH.2
PKMA
PARP6a
SI:DKEY-205H23.1
FAM219B
Dre 18*
HCN4
DHDH
PKMA
PARP6a
COX5AB
CABZ01078320.1
FAM219B
RPP25B
CELF6
Dre 25
HCN4L
PKMB
PARP6b
RPP25A
COX5AA

Figure. 37. *PARP6* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study.

The NCBI Genome and Ensembl Databases were used to identify *PARP6's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (Cunningham et al. 2022, NCBI 2022). Dre 18's *PARP6* genomic context originates from NCBI, while Dre 18's **PARP6* genomic context originates from Ensembl, both are provided at full disclosure as there are surrounding genes that appear similar to other listed species. The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP6* gene is not present in *Drosophila*.

The surrounding context relative to *PARP7* in mouse, rat, and zebrafish that has maintained order relative to human includes *KCNAB1* (*KCNAB1A* in zebrafish), *SSR3,* and *LEKR1* (Fig. 38, NCBI, Tables 7-11). *KCNAB1*, *SSR3*, and *LEKR1*, and *CCNL1* have maintained order in human, mouse, and rat. Zebrafish's context has maintained gene order overtime except for the presence of the *LOC101885893* gene. In zebrafish *CCNL1A* is an ortholog to the human *CCNL1* gene, and is located on chromosome 18, but is further down the chromosome from the listed surrounding genes. *CCNL1B* is also an ortholog to human *CCNL1* located on chromosome 2.

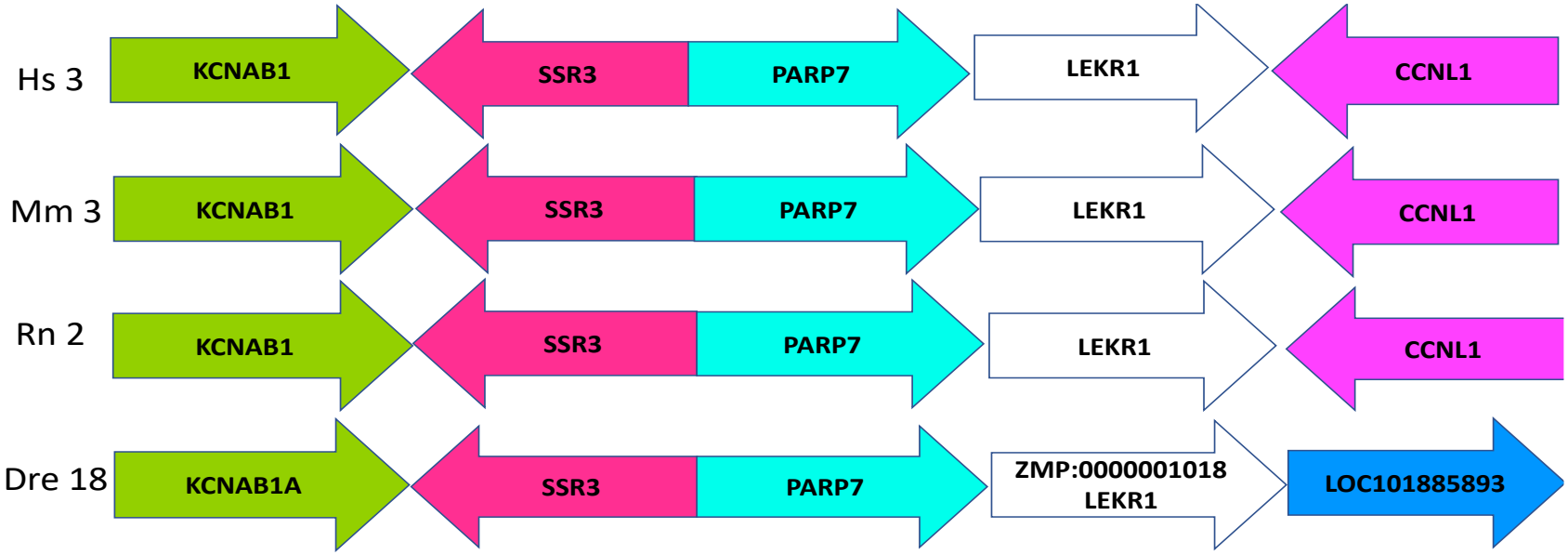


Figure. 38. *PARP7* Chromosomal Context Shows Similar Conserved Gene Order for Four Out of Five Species in this Study.

The NCBI Genome Database was used to identify *PARP7's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* genes were used as the center reference when generating each of the other PARP species gene order contexts. The *PARP7* gene is not present in *Drosophila*.

Overtime, relative to *PARP8*'*s* context, *HCN1*, *EMB*, *PARP8*, *ISL1*, *ITAG1*, and *PELO* have maintained the same order across mouse and rat relative to human (Fig. 39, NCBI, Tables 7-11). Note that *ISL1*, *ITAG1*, and *PELO* genes in human is further down on chromosome 5. *ISL1*, *ITAG1*, and *PELO* in mouse is further up on chromosome 13 and are also further up chromosome 2 in rat. In zebrafish one of the surrounding genes is *ISL1A*, which is like the *ISL1* gene in the other noted species and is located on chromosome 5. There is an *IS1LB* gene that is located on chromosome 10. *HCN1* is not present in zebrafish, *EMB* is located on chromosome 5, and *ITGA1* is located on chromosome 10.

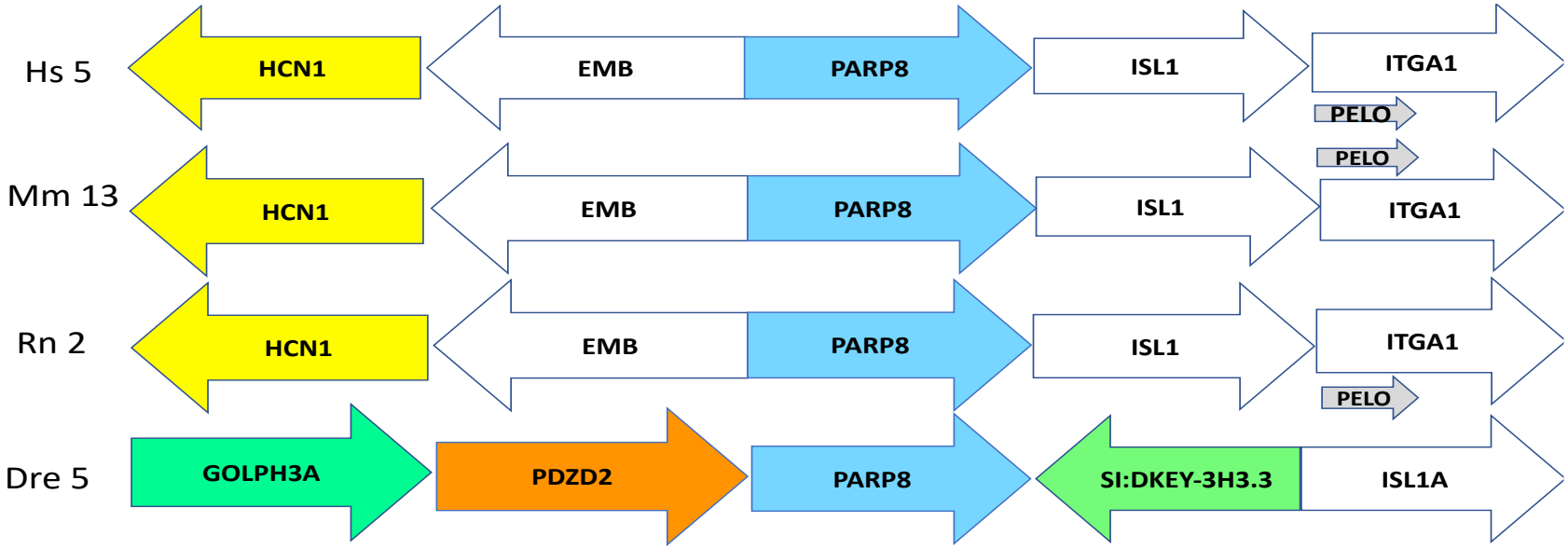


Figure. 39. *PARP8* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study.

The NCBI Genome Database was used to identify *PARP8's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP8* gene is not present in *Drosophila*.

PARP9 and *PARP14* reside near each other on chromosome 14 in human, the same genes, and similar context was also identified across mouse, rat, and zebrafish species (Fig. 40, NCBI, Tables 7-11). *PARP15*, which is near human *PARP14*, is not present in mouse, rat, or zebrafish. Relative to human, mouse and rat have maintained context order. In zebrafish, the order is somewhat similar, but those surrounding genes are either found on other chromosomes or they are not present. In zebrafish *HSPBAP1*, *SLC49A4,* and *PARP14rs3* are located further down chromosome 9. The *DTX3L* gene in human is an ortholog to the *DTX3LB.2*, *DTX3LB.3*, and *DTEX3LB.1* genes in this species and these are located on chromosome 11. *KPNA1* is located on chromosome 24 and *WDR5B* is not present in this species. *ACTR3*, *GPR39*, and *UBXN4* are the genes that surround *PARP14rs3* (*PARP14_2*). *PARP14rs4* (*PARP14_3*) is located on chromosome 10 but appears to be more closely related to *PARP14rs1* (*PARP14_1*) according to phylogenetic analysis (Fig. 20, NCBI 2022). In synteny analysis *PARP9* and *PARP14* are present on both human and mouse chromosomes, and that the lack of *PARP15* in mouse is not due to a chromosomal rearrangement, as they reside in single syntenic genomic blocks (Fig. 41, JAX Synteny Browser Kolishovski et al. 2019).

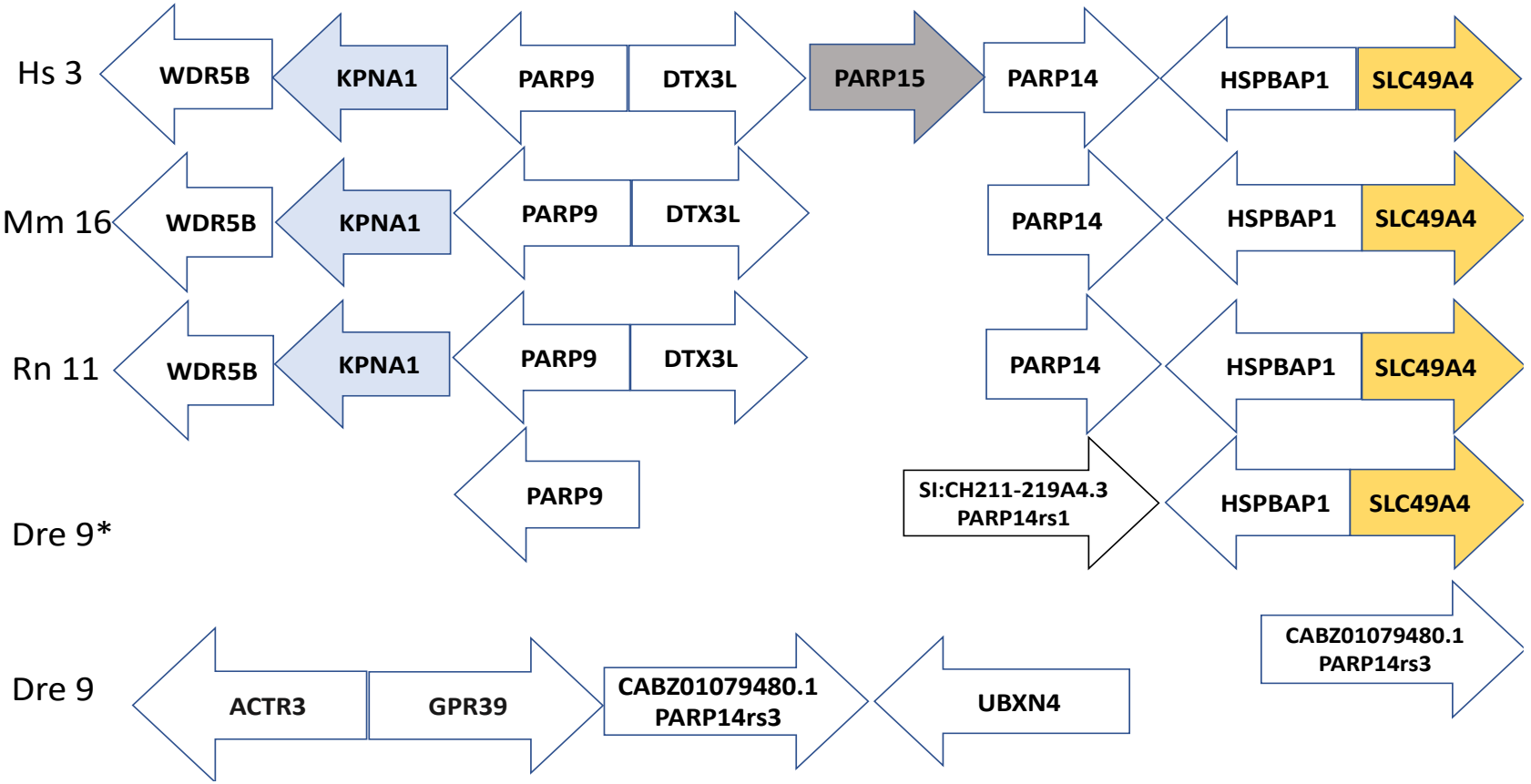

Hs 3
WDR5B
KPNA1
PARP9
DTX3L
PARP15
PARP14
HSPBAP1
SLC49A4
Mm 16
WDR5B
KPNA1
PARP9
DTX3L
PARP14
HSPBAP1
SLC49A4
Rn 11
WDR5B
KPNA1
PARP9
DTX3L
PARP14
HSPBAP1
SLC49A4
PARP9
SI:CH211-219A4.3
PARP14rs1
HSPBAP1
SLC49A4
CABZ01079480.1
PARP14rs3
Dre 9*
Dre 9
ACTR3
GPR39
CABZ01079480.1
PARP14rs3
UBXN4

Figure. 40. *PARP9*, *PARP14*, and *PARP15* Chromosomal Context Shows Similar Conserved Gene Order for Four Out of Five Species in this Study. The NCBI Genome and Ensembl Databases were used to identify *PARP9's*, *PARP14*'s, and *PARP15's* location within each of listed species genomes and the direction of transcription. *PARP9, PARP14_1* (Dre-9*, *SI:CH211-219A4.3* was previously listed on Ensembl but is no longer present and still appears on the zebrafish information network-ZFIN database) appears to be same as *PARP14rs1* (NCBI), and *PARP14_2* (*CABZ01079480.1*-Ensembl) appears to be same as *PARP14rs3* (NCBI) and has been identified in Dre (Bradford et al. 2022, Cunningham et al. 2022, NCBI 2022). Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (Cunningham et al. 2022, NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The gene context that surrounds *CABZ01079480.1* is listed in zebrafish as it appears to be closely related to *PARP14* and is known as a human *PARP15* ortholog (Cunningham et al. 2022). *PARP14rs2.1* was not chosen for analysis. *PARP9*, *PARP14*, and *PARP15* are absent in *Drosophila*.

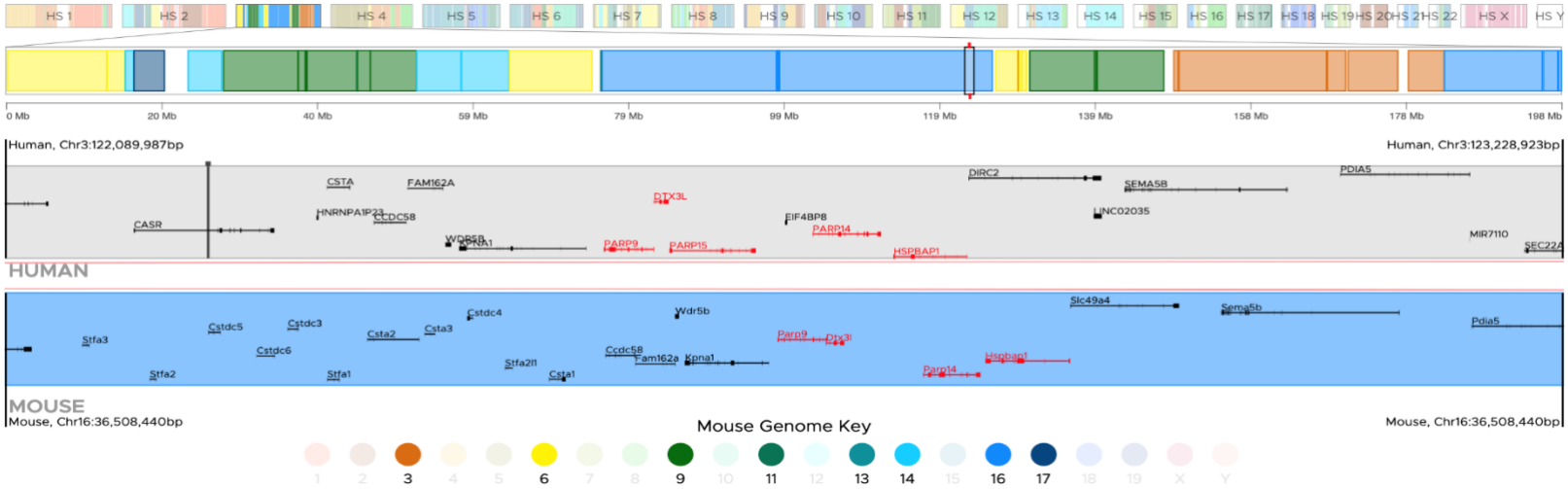


Figure. 41. Human-Mouse Synteny surrounding *PARP9*, *PARP14*, and *PARP15*. Gene order and changes visualized in the JAX Synteny Browser (Kolishovski et al. 2019) with A) Human compared to Mouse. Genes highlighted in red are *PARP9*, *PARP14*, *PARP15*, *DTX3L*, and *HSPBAP1*.

In *PARP10*, similar context order has been maintained in human, mouse, and rat species (Fig. 42, NCBI, Tables 7-11). The location of *GRINA* was observed within *PARP10* in humans, but was located upstream of *PARP10* in mouse and rat. There appears to be an annotation error on NCBI because human genomic context from Ensembl was compared to that of the human and mouse context from the JAX Synteny Browser (Fig. 43 A, B). Therefore, chromosome context is considered identical in human, mouse, and rat. In zebrafish the same surrounding genes cannot be observed as *SPAG1A*, *POLR2K*, *CPSF1*, and *ADCK5* are present. In zebrafish the *SPATC1* gene is not present. *GRINAA* is an ortholog to the human *GRINA* gene and is located on chromosome 19. *PLECA* is an ortholog to the human *PLEC* gene and is located on chromosome 19, while *PLECB* is located on chromosome 16. *EPPK1* is located on chromosome 19.

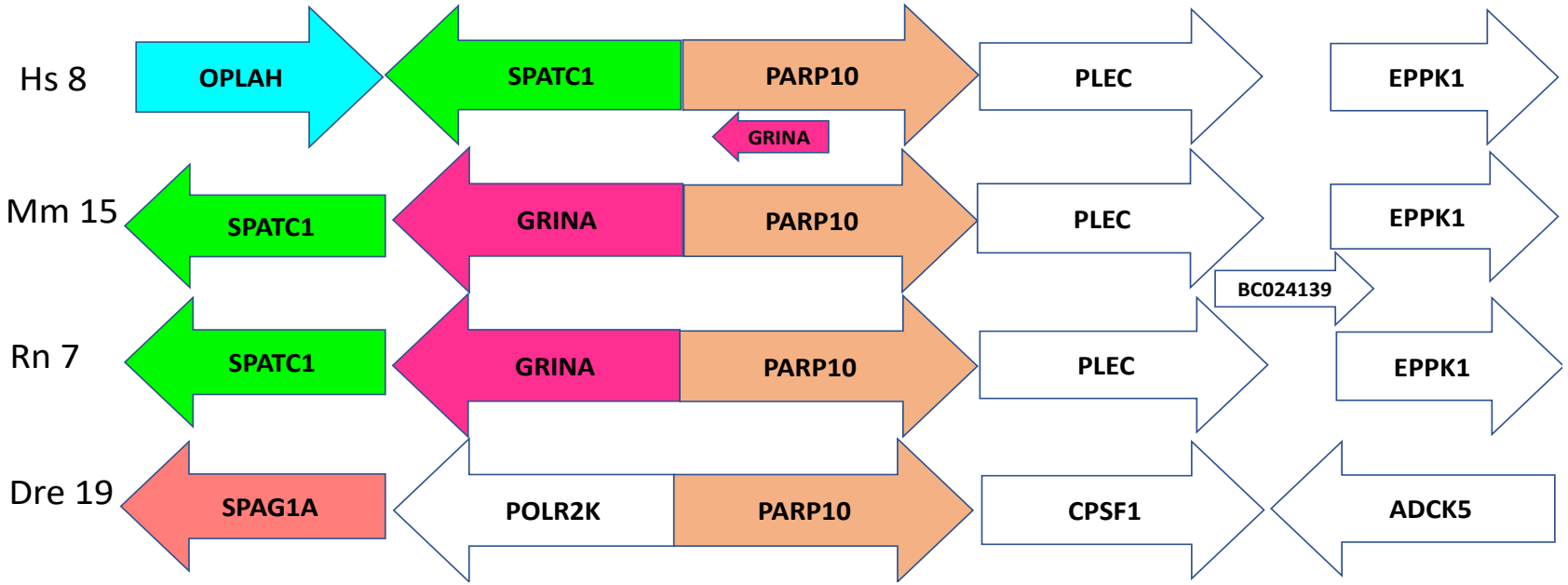


Figure. 42. *PARP10* Chromosomal Context Shows Similar Conserved Gene Order for Three Out of Five Species in this Study. The NCBI Genome Database was used to identify *PARP10's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP10* gene is not present in *Drosophila*.

A

B

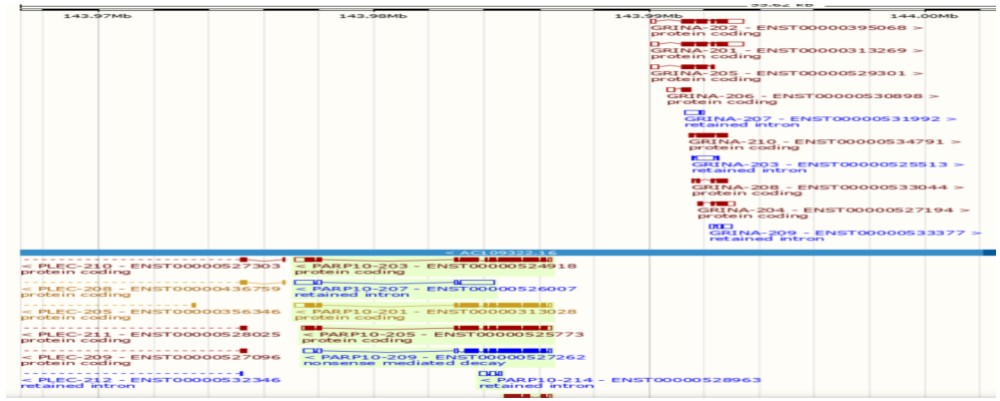


Figure. 43. Human-Mouse Synteny surrounding *PARP10* alongside *PARP10's* Ensembl Genomic Context. Gene order and changes visualized in the JAX Synteny Browser (Kolishovski et al. 2019) and Ensembl (Cunningham et al. 2022) with A) Human compared to Mouse and B) Human. Genes highlighted in red are *PLEC*, *GRINA*, *PARP10*, and *SPATC1* (A). The same genes are noted in (B).

For *PARP11*, surrounding genes are found to be similar in human, mouse, and rat (Fig. 44, NCBI and Ensembl, Tables 7-11). The genes that surround *PARP11* in zebrafish are *FTR64*, *PYROXD1*, *IAPP*, and *SLCO1F1*. In zebrafish *TIGARA* is an ortholog to the human *TIGAR* gene and is located on chromosome 25. *TIGARB*, which is also an ortholog, is located on chromosome 4. *CCND2A*, which is an ortholog to the human gene *CCND2* is located on chromosome 25. Similarly, *CCND2B* is located on chromosome 4. *CRACR2AA* is an ortholog to the human *CRACR2A* and is located on chromosome 18. *CRACR2AB* is also located on chromosome 4. *PRMT8B* is an ortholog to the human *PRMT8* gene and is located on chromosome 4.

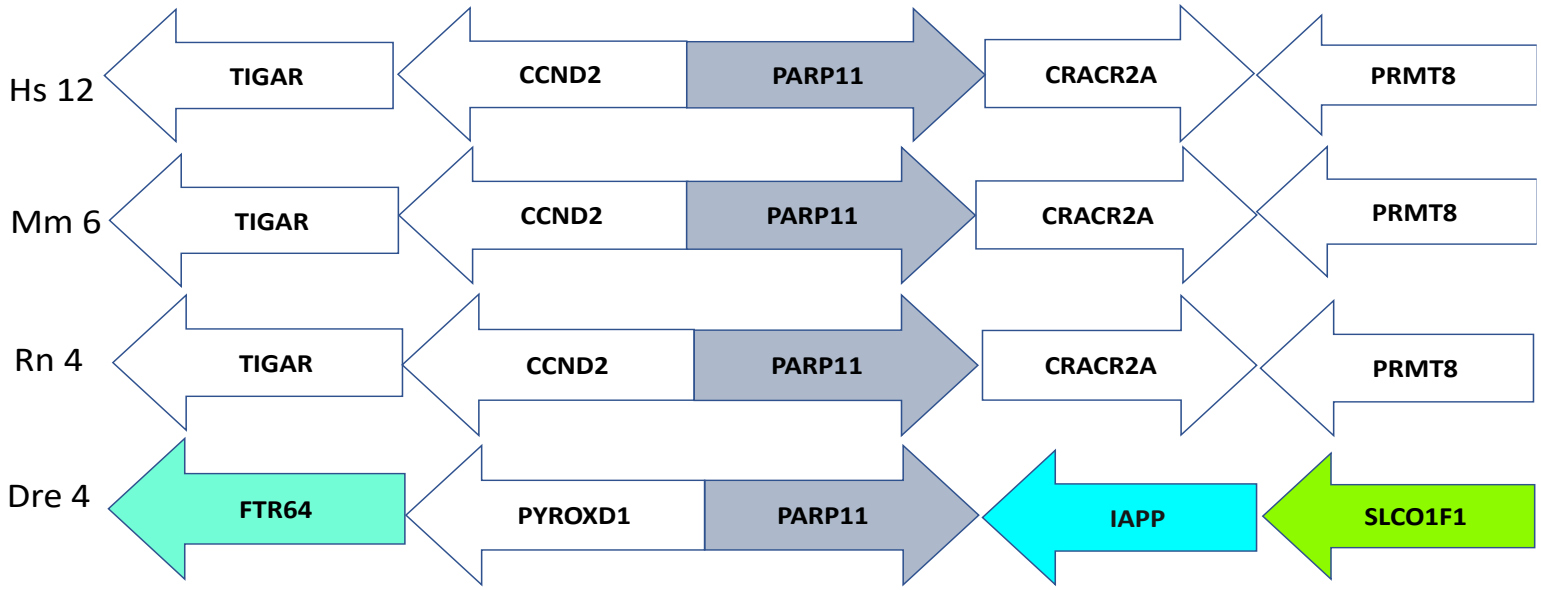


Figure. 44. *PARP11* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study. The NCBI Genome and Ensembl Databases were used to identify *PARP11's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (Cunningham et al. 2022, NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. The *PARP11* gene is not present in *Drosophila*.

Chromosome context relative to *PARP12* and *PARP13* has been maintained in human, mouse, and rat (Fig. 45, NCBI, Tables 7-11). Upstream of *PARP12a* in zebrafish, *TBXAS1*, *HIPK2*, and *KIAA1549* can be found, indicating that its context has maintained some order when compared to the other species. Note that the *KIAA1549* gene in human has orthologs in all listed species with the exception of *Drosophila* and are listed as follows: mouse *D630045J12RiK*, rat *RGD1306271*, zebrafish (12a) *SI:CH211-1E14.1*, and zebrafish (12b) *SI:DKEYP-27E10.3* these genes were listed as *KIAA1549* for consistency across species genomic context. In zebrafish *SLC37A3* is located on chromosome NW_003336877.1. *KDM7AA* which is an ortholog to the human *KDM7A* gene is located on chromosome 4. *KDM7AB* is located on chromosome 25. *UBN2A* which is an ortholog to human gene *UBN2* is located on chromosome 6. *UBN2B* is located on chromosome 3 and *TTC26* is located on chromosome 6.

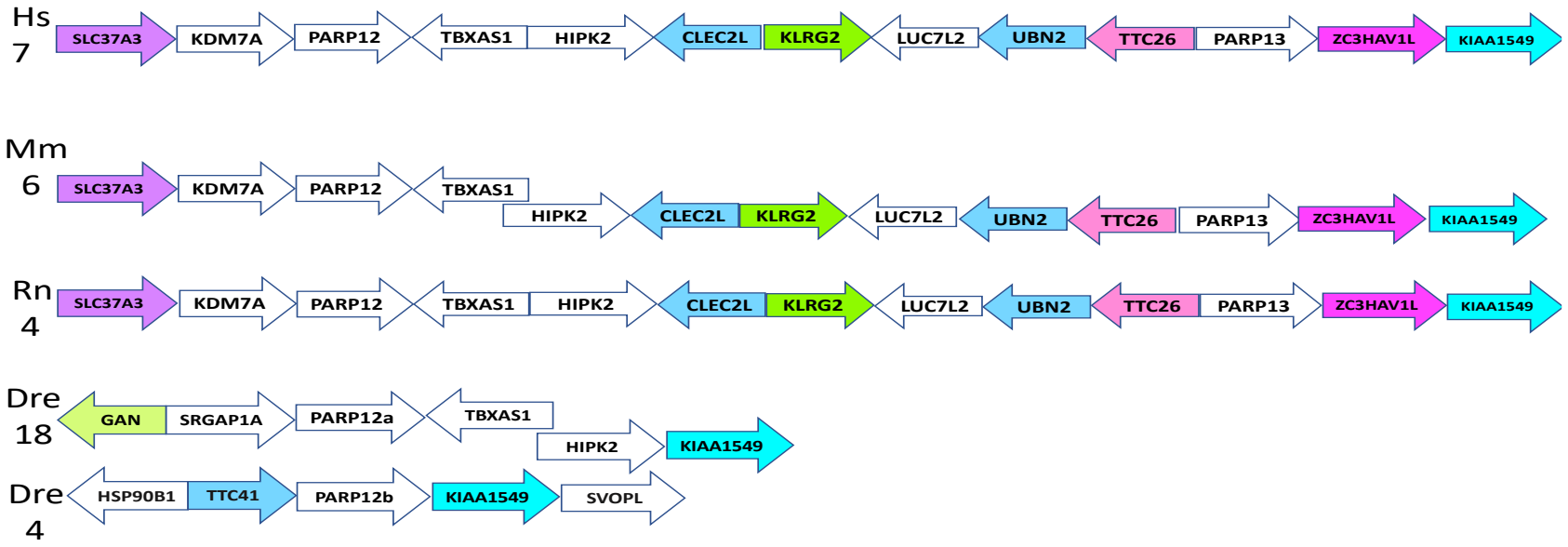


Figure. 45. *PARP12* and *PARP13* Chromosomal Context Shows Similar Conserved Gene Order for Four Out of Five Species in this Study. The NCBI Genome Database was used to identify *PARP12's* and *PARP13's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order contexts. *PARP12* and *PARP13* are not present in *Drosophila*.

Interestingly, *PARP13* is also known as *ZC3HAV1* and a gene named *ZC3HAV1L* resides immediately downstream, so we investigated their similarity. Synteny analysis shows that on the human chromosome 7 and mouse chromosome 6 that *PARP13* (*ZC3HAV1*) and *ZC3HAV1L* have a preserved genomic context (direct repeats), as well as *PARP12* and other surrounding genes *TBXAS1*, *HIPK2*, *KDM7A*, *UBN2*, *TTC26*, *LUC7L2*, *CLEC2L*, and *KLRG2* (Fig. 46, JAX Synteny Browser Kolishovski et al. 2019). *PARP13* and *ZC3HAV1L* have similar percent identity when their amino acid sequences are aligned (Fig. 47 A). However, *ZC3HAV1L* appears to be truncated at its carboxy-terminal end compared to *PARP13* (Fig. 47 B). To place *ZC3HAV1L* within the gene family phylogenetic analysis with all species was repeated including this gene. As expected, *ZC3HAV1L* in human, mouse, and rat grouped closely to *PARP13* and also a CCCH-PARP (Fig. 48).

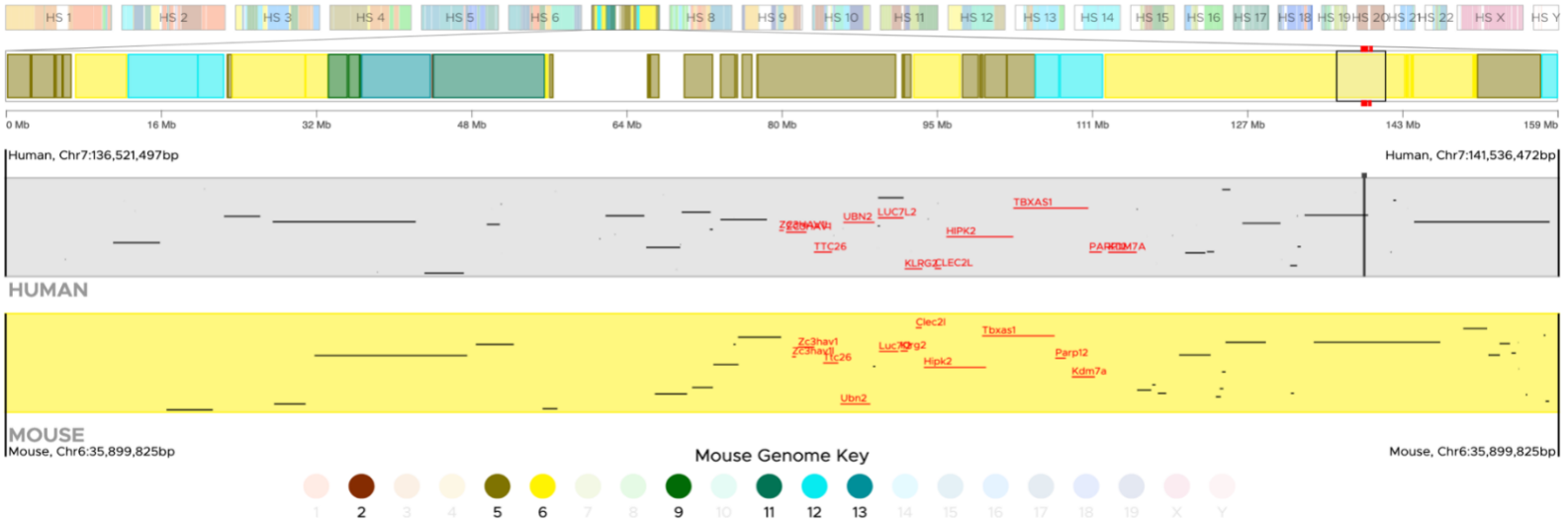
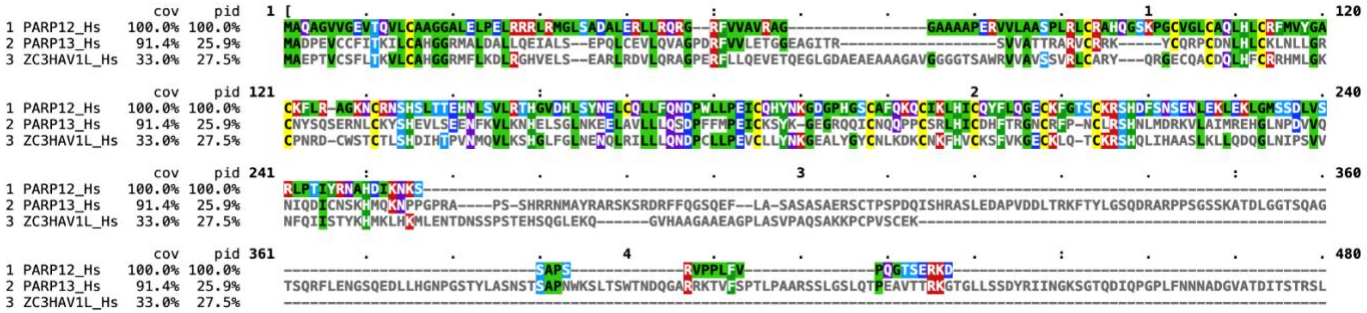


Figure. 46. Human-Mouse Synteny surrounding *PARP13* (*ZC3HAV1*). Gene order and changes visualized in the JAX Synteny Browser (Kolishovski et al. 2019) with A) Human compared to Mouse. Genes highlighted in red are *PARP12*, *PARP13*, *ZC3HAV1L*, *TBXAS1*, *HIPK2*, *KDM7A*, *UBN2*, *TTC26*, *LUC7L2*, *CLEC2L*, and *KLRG2*.

A



B

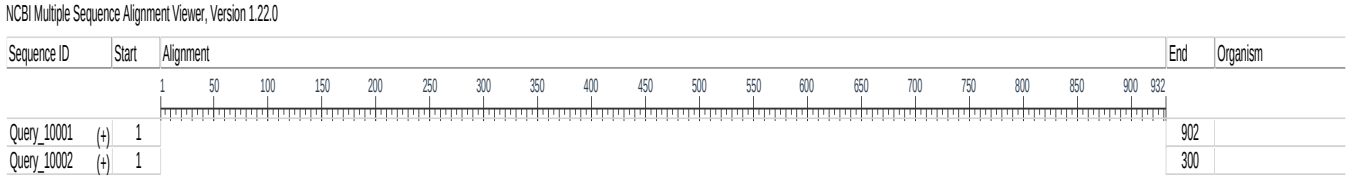


Figure. 47. *ZC3HAV1L* is similar to *PARP13* according to Clustal Omega-MView and COBALT Multiple Sequence Alignments. In panel A, protein sequence alignment of human proteins PARP12, PARP13, and ZC3HAV1L was done with the use of Clustal Omega and MView multiple alignment tools (Kent et al. 2002, Cunningham et al. 2022, Sievers et al. 2011, Brown et al. 1998, Brown 1999; Table 3). In panel B, COBALT multiple sequence alignment was done with human PARP13 (Query_10001) and ZC3HAV1L (Query_10002) proteins (NCBI 2022, Papadopoulos and Agarwala 2007, Kent et al. 2002, Cunningham et al. 2022; Table 3).

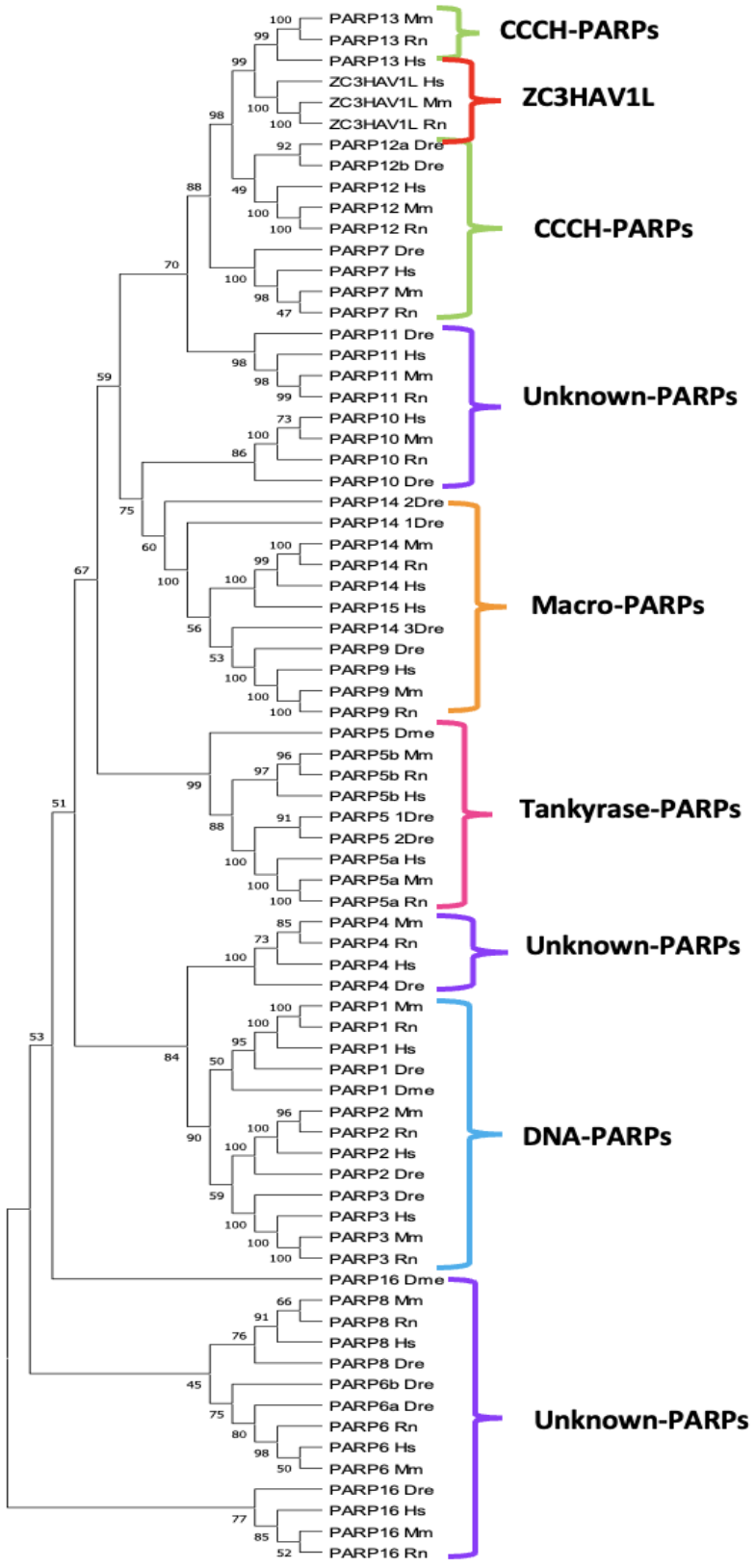

PARP13 Mm
PARP13 Rn
PARP13 Hs
ZC3HAV1L Hs
ZC3HAV1L Mm
ZC3HAV1L Rn
PARP12a Dre
PARP12b Dre
PARP12 Hs
PARP12 Mm
PARP12 Rn
PARP7 Dre
PARP7 Hs
PARP7 Mm
PARP7 Rn
PARP11 Dre
PARP11 Hs
PARP11 Mm
PARP11 Rn
PARP10 Hs
PARP10 Mm
PARP10 Rn
PARP10 Dre
PARP14 2Dre
PARP14 1Dre
PARP14 Mm
PARP14 Rn
PARP14 Hs
PARP15 Hs
PARP14 3Dre
PARP9 Dre
PARP9 Hs
PARP9 Mm
PARP9 Rn
PARP5 Dme
PARP5b Mm
PARP5b Rn
PARP5b Hs
PARP5 1Dre
PARP5 2Dre
PARP5a Hs
PARP5a Mm
PARP5a Rn
PARP4 Mm
PARP4 Rn
PARP4 Hs
PARP4 Dre
PARP1 Mm
PARP1 Rn
PARP1 Hs
PARP1 Dre
PARP1 Dme
PARP2 Mm
PARP2 Rn
PARP2 Hs
PARP2 Dre
PARP3 Dre
PARP3 Hs
PARP3 Mm
PARP3 Rn
PARP16 Dme
PARP8 Mm
PARP8 Rn
PARP8 Hs
PARP8 Dre
PARP6b Dre
PARP6a Dre
PARP6 Rn
PARP6 Hs
PARP6 Mm
PARP16 Dre
PARP16 Hs
PARP16 Mm
PARP16 Rn
CCCH-PARPs
ZC3HAV1L
CCCH-PARPs
Unknown-PARPs
Macro-PARPs
Tankyrase-PARPs
Unknown-PARPs
DNA-PARPs
Unknown-PARPs

Figure. 48. PARP Species Phylogenetic Analysis with *ZC3HAV1L*. ZC3HAV1L's protein sequence is analyzed alongside all PARP species protein sequences where the Maximum Likely hood method-JTT matrix-based model was used and applied to 74 sequences with 100 tree bootstrapping replicates and 3115 amino acid positions (Jones et al. 1992, Felsenstein 1985; Tables 3-6). Each of the PARP groups are labeled and colored according to function. Green indicates the CCCH-PARPs, purple indicates the unknown (undefined)-PARPs, orange indicates the Macro-PARPs, pink indicates the Tankyrase-PARPs, red indicates ZC3HAV1L, and blue indicates the DNA-PARPs.

PARP16 has maintained identical surrounding gene context in human, mouse, and rat species (Fig. 49, NCBI, Tables 7-11). Zebrafish and fruit fly have dissimilar context. In zebrafish *IGDCC4* is located on chromosome 7, *IGDCC3* is located on chromosome 7, *CILP* is located on chromosome 7, and *CLPXA* which is an ortholog to the human gene *CLPX* is also located on chromosome 7, while *CLPXB* is located on chromosome 25. *IGDCC4*, *IGDCC3*, *CILP*, and *CLPX* are not present in fruit flies. There is a *clpX* gene in fruit flies but has a different function then the gene listed in this chromosome context in the other species (NCBI 2022).

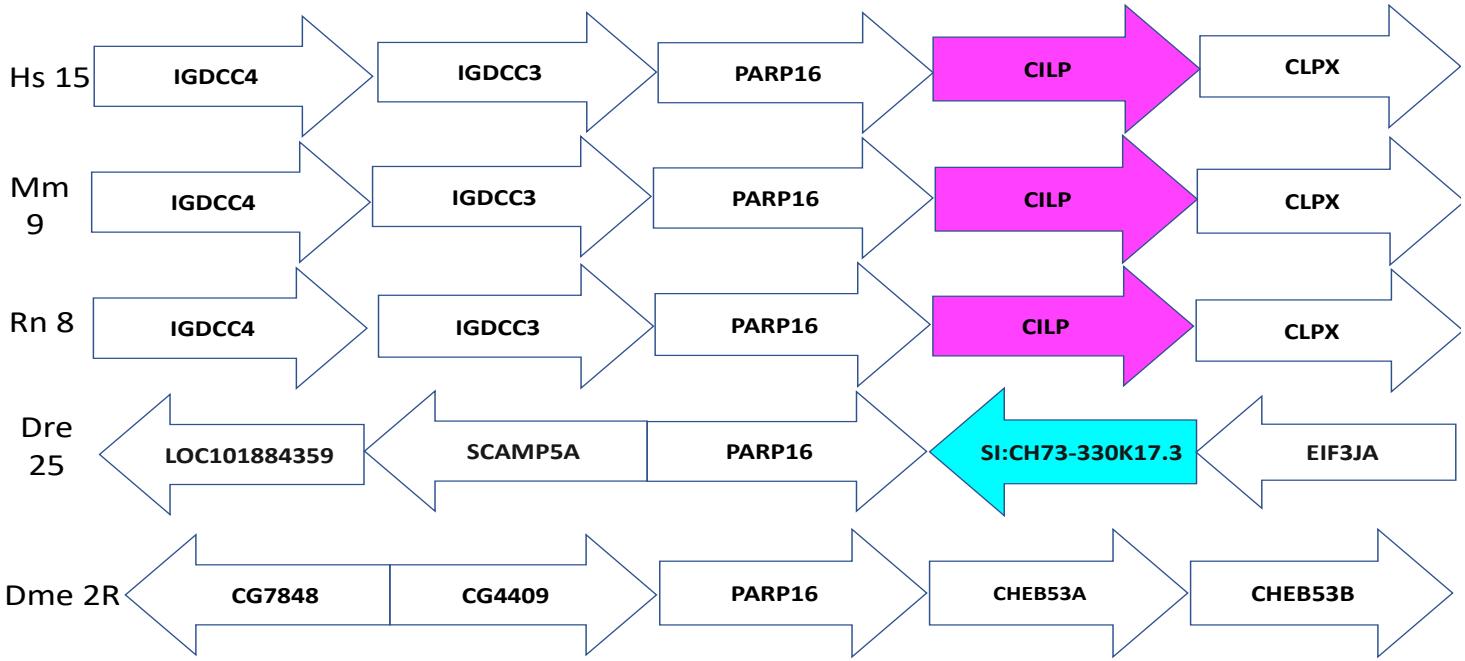


Figure. 49. *PARP16* Chromosomal Context Shows Conserved Gene Order for Three Out of Five Species in this Study. The NCBI Genome Database was used to identify *PARP16's* location within each of listed species genomes and the direction of transcription. Homologs are the same color and relative direction of transcription is indicated by the arrow direction. The numbers listed by the species abbreviations are the chromosome numbers that indicate where genes are located (NCBI 2022). The human *PARP* gene was used as the center reference when generating each of the other PARP species gene order context.

DISCUSSION

The goals of this project were to 1) identify relationships among *PARP* gene family members and 2) identify the undefined-PARP's nearest relative in the gene family using multiple evolutionary lines of evidence. In previous gene family studies conducted by Ocampo Daza et al. (2011) and Tostivint et al. (2014), similar analyses done in this study for the *PARP* gene family were also used to analyze the Insulin-Like Growth Factor Binding Protein (*IGFBP*) gene family members and G-protein-coupled receptors (*SSTR* and *UTS2R*) gene family members (specifically receptors of somatostatin (SS) and urotensin II (UII)) across species, including sequence alignments, phylogenetic, and genomic context analyses to analyze the evolution of members and evolutionary events, such as ancestral gene members, variances in functions, as well as gene duplications.

Amino Acid Analyses

First, by looking at amino acid similarity and identity percentages in PARP species pairs we found that human, mouse, rat, and zebrafish family members generally grouped by function, but percentages were generally low, indicating this is likely an older gene family. Multiple alignment per species showed moderate similarity, which mirrors the long-ago divergence. Some interesting features within these data are that DNA-PARP members and Tankyrase-PARP members across all species had amino acid identity and similarity percentage ranges that supported their relations as groups. Unknown-PARPs 6 and 8 had higher amino acid identity and similarity to each other regardless of species. Finally, a large group with some shared identity and similarity are

PARPs 9-15 and PARP7, which are mostly CCCH and Macro-PARPs. The conclusions that can be best supported from these analyses is that PARP10 is more closely related to Macro-PARPs, PARP11 is more closely related to CCCH-PARPs, PARP4, PARP6, PARP8, and PARP16 are most related to DNA-PARPs. COBALT analyses indicate that these family members retain some close relations but have also diverged as conservation of amino acids are not relatively high.

Phylogenetic Analyses

Phylogenetic analyses supported that known functional group PARPs are related with one another as well as mostly supported the same relationships of the undefined PARPs that were found with the amino acid similarity and identity analyses. PARP11 is most related to CCCH-PARPs as indicated by its relationship with PARP12 and PARP13. PARP10 is related to both or either Macro or CCCH-PARPs, but more support was found for relatedness to Macro-PARPs. PARPs 4, 6, 8, and 16 appear to be more closely related with DNA-PARPs. Undefined-PARP11 grouped with the CCCH-PARPs. PARP14_2 in zebrafish is likely an ortholog of human PARP10, as it grouped with PARP10 rather than PARP14.

Expression Analyses

Expression analyses provided insight on how gene expression across all species has been conserved overtime. Pairwise co-expression of human *PARP9* and *PARP14* is consistent with Macro-PARP protein grouped functions. The co-expression observed between *PARP9* and *PARP12* and *PARP12* and *PARP14* is interesting in that *PARP12* is a CCCH-PARP and *PARP14* is a Macro-PARP. *PARP10* appeared to co-express with *PARP13*,

another CCCH-PARP, and the other mentioned PARPs to a lesser extent, suggesting that *PARP10* is a CCCH-PARP or Macro-PARP. DNA-PARPs (1, 2) and Tankyrase-PARPs (5a, 5b) showed some co-expression indicating consistency with grouped functions. In mouse, *PARP9* and *PARP14*'*s* co-expression is also consistent with Macro-PARP grouped functions, and the moderate co-expression observed between *PARP9* and *PARP12* and *PARP12* and *PARP14* was interesting as this was similarly observed in humans. *PARP10*'s co-expression with *PARP9* indicates it maybe most related to Macro-PARPs. *PARP11* and *PARP13* as well as *PARP3* and *PARP4* in mouse were similarly co-expressed as well as with these groups mentioned above indicating *PARP11* could be a CCCH-PARP and *PARP4*, which further supports it is a relative of DNA-PARPs. DNA-PARPs (1, 2) and Tankyrase-PARPs (5a, 5b) shared some co-expression, suggesting a relationship.

The clearest identified co-expressed relationship occurs between the Macro-PARPs (9 and 14) in rat, indicating that the Macro-PARPs have the strongest relationship relative to expression in this analysis, as human, mouse, and zebrafish results were similar. The moderate pairwise co-expression relationship of *PARPs* 9 and 12, as well as 12 and 14 were observed, as similarly viewed in mouse and humans. *PARP10* and *PARP13* join these groups in co-expression, again suggesting *PARP10* may be a CCCH or Macro-PARP member. *PARP3* and *PARP4* also showed some co-expression with the listed groups above, again supporting *PARP4* as a DNA-PARP relative. Tankyrase-PARPs, 5a and 5b are consistent with their grouping as they are somewhat co-expressed with one another. The clearest identified co-expressed relationship occurs between the Macro-PARPs (9 and 14_1) in zebrafish, indicating that the Macro-PARPs have the

strongest PARP relationship with respect to gene expression in this analysis, as human, mouse and rat results were similar. *PARP10*, *PARP12a*, and *PARP14_3* also share co-expression amongst one another and those listed above, suggesting *PARP10* may either be a CCCH or Macro-PARP.

Group cluster analysis of gene expression in normal human tissue samples align with grouping of protein functions as samples were broadly grouped into three main groups according to expression. *PARP15* is the single member of group A, which indicates high divergence or a specialized function as its expression was confined to just a few tissues. Group B has mostly Macro or CCCH-PARPs as well as undefined-PARPs 4, 8, and 10, suggesting similarity to these groups. Group C had mainly DNA, CCCH, and Tankyrase-PARPs as well as undefined-PARPs 11, 6, and 16. However, the inclusion of most PARP function groups in group C make any conclusion very speculative. The best supported conclusions from gene expression analysis gives further support for the results previously obtained in the above analyses, that *PARP10* is likely related to Macro-PARPs or CCCH-PARPs, *PARP11* is most similar to CCCH-PARPs, and that *PARP4*, *PARP6*, *PARP8*, and *PARP16* are most similar to DNA-PARPs.

Gene Order Analyses

Gene order context identified genome conservation or divergence events that occurred relative to each of the studied species. Gene order context relative to *PARP1*, *PARP2,* and *PARP3* has been relatively conserved across species. *PARP4*'s gene order context in human compared to mouse and rat indicates that a chromosomal inversion

event occurred as the altered orientation of genes *ATP12A* and *RNF17* is present. The *CENPJ* gene is also located in a different region on the chromosome in human, which indicates an event happened as this gene's location in mouse and rat has been conserved overtime. Synteny genomic analysis in human confirmed that both *ATP12A* and *RNF17* represent a chromosomal inversion event, as well as *PARP4* itself. The *TNKSa* gene in zebrafish is the most closely related zebrafish member to the *PARP5a* gene in human, mouse, and rat according to gene order. Synteny analysis confirms that *DUSP4* is not conserved in the same chromosomal region relative to both human and mouse genomic context. Gene context in zebrafish on chromosome 5 (*TNKSb*) and in fruit fly differs from human, mouse, and rat indicating that events occurred to alter the similar context that has been preserved overtime in these species. The only context that differs from the other species is fruit fly relative to *PARP5b*. The *TNKS* gene in fruit fly has been identified as an ortholog to both human *PARP5a* and *PARP5b* genes. *PARP6*, *PARP7*, and *PARP8* gene context has been relatively conserved across species. *PARP9*, *14*, and *15*'*s* genomic context indicates that *PARP15* is likely a recent gene duplication event from *PARP14* in human, as *PARP15* is also present in chimpanzees (data not shown), and is not present in the other listed species. *PARP15* may also have acquired some sort of specialized function or expression in human due to its narrow tissue-specific expression profile. Indeed, a prior analysis by Perina et al. (2014) suggested that PARP14 was often duplicated, and duplicates were found predominantly in vertebrates. They furthermore suggested that PARP15 is likely a duplication of PARP14 and that this originally occurred before mammals split from the common ancestor (Perina et al. 2014). Evidence

supporting this duplication is several-fold. First, the Macro-PARP domains in PARP9 and 14 are relatively similar in length but PARP15's macro domain appears to be shorter (Li and Chen 2014). *PARP15* is the family member that shows the lowest expression and is absent in all species except humans; since it is present in chimps (data not shown), we suggest it also may be a single duplication of *PARP14* in humans according to our analyses, that it may have acquired very specific expression or a specialized function that the other species do not have, or that it may be undergoing pseudogenization or could have been deleted in other species. The human protein atlas contained RNA-seq tissue data (not shown) for both *PARP14* and *PARP15* and revealed that both genes are expressed in the bone marrow/lymphoid and gastrointestinal tissues but that *PARP15* is expressed at a lesser extent (Uhlén et al. 2015). Gene cluster expression data from tissues (not shown) showed that *PARP14* is present amongst other genes in clusters of macrophages that carry out immune response functions, while *PARP15* is present amongst other genes in clusters that are specific to general immune cell response functions, which again suggests that *PARP15* may have acquired specific expression and function (Uhlén et al. 2015). Synteny genomic analysis indicates that *PARP9*, *PARP14*, and *PARP15* in human are likely all duplications of one another, as they reside near each other on the same chromosome. Similarly, *PARP11's*, *PARP12's*, *PARP13's*, and *PARP16's* genomic context has been relatively conserved across species.

ZC3HAV1L

In the study of chromosome context, we noticed a possible new *PARP* gene family member. Genomic context-synteny analyses confirms that *Z3CHAV1L* is a likely gene duplication of *PARP13* which resulted from a tandem repeat. According to sequence analyses, human *ZC3HAV1L* is more closely related to CCCH-PARPs 12 and 13 as confirmed via phylogenetic tree analysis.

General Conclusion

Taken together, human Macro-PARPs (9-14) have retained the strongest group relation overtime, as function, phylogenetic, gene expression, and gene order analyses agree to these relationships. Human DNA-PARPs and Tankyrase-PARPs may also have been recently duplicated as they retain higher pairwise amino acid identity-similarities, close relations according to phylogenetic analysis, grouped expression relative to pairwise co-expression and RNA-seq human tissue analysis, and gene order context that has been relatively conserved. CCCH-PARPs across species are similar relative to pairwise amino acid identity-similarities, phylogenetic analysis (humans, mouse, and rats), gene expression, and gene order context analyses (humans, mouse, rats, and zebrafish). Undefined-PARPs, PARP4, PARP6, PARP8, and PARP16 are likely related to DNA-PARP functional group based on pairwise amino acid identity-similarity, phylogenetic, and gene expression analyses. PARP10 is likely related to CCCH, or Macro-PARP functional group based on pairwise amino acid identity-similarity, phylogenetic, and gene expression analyses in species. PARP11 is most likely related to the CCCH-PARP functional group based on pairwise amino acid identity-similarity, phylogenetic, and

gene expression analyses. PARP11 was also suggested to have lost Zinc activity due from a duplication event, which aligns with our study in that PARP11 could be a part of the CCCH-Zinc PARP family (Perina et al. 2014).

When we further compared the results of our current study to prior studies (Citarelli et al. 2010, Perina et al. 2014), many overlaps can be found. In these prior studies of eukaryotic PARPs, family members were organized according to protein related clade groupings. Clade 1 incorporated human PARP members 1, 2, and 3 which is consistent with DNA-PARPs in our analyses and PARP1 was identified as the likely ancestral PARP gene member (Citarelli et al. 2010, Perina et al. 2014). Clade 1 was also found in mouse, zebrafish, and fruit fly species (PARP1) (Perina et al. 2014). Clade 2 was not found in our species models. Clade 3 included human members 9, 14, 15, 7, 12, 13, 10, and 11 (Citarelli et al. 2010, Perina et al. 2014) which is consistent with our findings as PARP9, PARP14, and PARP15, are Macro-PARPs, PARP7, PARP12, and PARP13 are CCCH-PARPs, and PARP10 is either a Macro or CCCH-PARP, and PARP11 is likely a CCCH-PARP. Similarly, clade 3 was found in mouse (except PARP15) and zebrafish (except PARP13 and PARP15) species (Perina et al. 2014). Clade 4 involved human, mouse, zebrafish, and fruit fly Tankyrase members PARP5a and PARP5b (except PARP5b in zebrafish) (Citarelli et al. 2010, Perina et al. 2014) which aligns with our results, as these members appear to be closely related across analyses. Clade 5 involved human, mouse, and zebrafish PARP4, an undefined-PARP, and suggested DNA-PARP member in our analyses, while clade 6 incorporated human, mouse, and zebrafish PARPs 6, 8, and 16

undefined members, which are also suggested to be DNA-PARPs in our analyses, and appear to share relations in amino acid and phylogenetic analyses across species, which suggests these PARPs may even form their own related sub-group (Citarelli et al. 2010, Perina et al. 2014). Citarelli et al. (2010) concluded that one ancestral protein likely functioned in response to DNA-damage similarly seen in PARP1 (human) and that the second ancestral gene may not have had ribosylation activity, but evolved it at a later time.

Taken together, these studies provide a more comprehensive and updated view of the evolutionary history of this gene family with our current study providing evidence of the nearest PARP relatives to undefined-PARPs, identification of gene evolution events in the Macro-PARP family members, and proposal of a new PARP member, ZC3HAV1L. Future studies should involve gene structure analysis to evaluate the conservation of exon structures that would lend additional molecular evidence to the similarity studies conducted above. Additional studies could include *PARP* genes in other species (from the 200 PARP homologs) to provide an overall consensus relative to conservation or divergence across family members. It may also be of value to focus some of these analyses specifically on the PARP domain, as this domain is key to defining these members as a gene family, is important for catalytic activity, and their functional groupings have been conserved overtime.

REFERENCES

Ahel D, Horejsí Z, Wiechens N, Polo SE, Garcia-Wilson E, Ahel I, Flynn H, Skehel M, West SC, Jackson SP, et al. 2009. Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. Science [Internet]. [cited 2022 Nov] ;325(5945):1240-3. Available from: https://doi.org/10.1126/science.1177321

Aguiar RC, Takeyama K, He C, Kreinbrink K, Shipp MA. 2005. B-aggressive lymphoma family proteins have unique domains that modulate transcription and exhibit poly(ADP-ribose) polymerase activity. J Biol Chem [Internet]. [cited 2022 Nov]; 280(40):33756-65. Available from: https:// doi.org/10.1074/jbc.M505408200

Alliance of Genome Resources Consortium. 2019. The Alliance of Genome Resources: Building a Modern Data Ecosystem for Model Organism Databases. Genetics [Internet]. [cited 2022 Jun]; 213(4):1189–1196. Available from: https://doi.org/10.1534/genetics.119.302523

Andersson DI, Jerlström-Hultqvist J, Näsvall J. 2015. Evolution of New Functions De Novo and from Preexisting Genes. Cold Spr Harb Pers Biol [Internet]. [cited 2021 Jan]; 7:a017996. Available from: https://doi.org/10.1101/cshperspect.a017996

Atasheva S, Akhrymuk M, Frolova EI, Frolov I. 2012. New PARP gene with an anti-alphavirus function. J Virol [Internet]. [cited 2022 Oct]; 86(15):8147-60. Available from: https://doi.org/10.1128/JVI.00733-12

Beck C, Robert I, Reina-San-Martin B, Schreiber V, Dantzer F. 2014. Poly(ADP-ribose) polymerases in double-strand break repair: focus on PARP1, PARP2 and PARP3.

Exp Cell Res [Internet]. [cited 2022 Nov]; 329(1):18-25. Available from: https://doi.org/ 10.1016/j.yexcr.2014.07.003

Bethesda (MD): National Library of Medicine (US). 2004. National Center for Biotechnology Information (NCBI). Gene [Internet]. [cited 2021 May]. Available from: https://www.ncbi.nlm.nih.gov/gene/

Bethesda (MD): National Library of Medicine (US). 2004. National Center for Biotechnology Information (NCBI). COBALT [Internet]. [cited 2022 Oct]. Available from: https://www.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi

Bick MJ, Carroll JW, Gao G, Goff SP, Rice CM, MacDonald MR. 2003. Expression of the zinc-finger antiviral protein inhibits alphavirus replication. J Virol [Internet]. [cited 2022 Nov]; 77(21):11555-62. Available from: https://doi.org/10.1128/jvi.77.21.11555-11562.2003

Bradford YM, Van Slyke CE, Ruzicka L, Singer A, Eagle A, Fashena D, Howe DG, Frazer K, Martin R, Paddock H, et al. 2022. Zebrafish information network, the knowledgebase for Danio rerio research. Genetics [Internet]. [cited 2022 Oct]; 220(4):iyac016. Available from: https://doi.org/10.1093/genetics/iyac016

Braschi B, Denny P, Gray K, Jones T, Seal R, Tweedie S, Yates B, Bruford E. 2019. Genenames.org: the HGNC and VGNC resources in 2019. Nuc Acids Res [Internet]. [cited 2021 Mar]; 47(D1):D786-D792. Available from: https://doi.org/10.1093/nar/gky930

Brown NP, Leroy C, Sander C. 1998. MView: A Web compatible database search or multiple alignment viewer. Bio info [Internet]. [cited 2022 Jun]; 14(4):380-381 Available from: https://doi.org/10.1093/bioinformatics/14.4.380

Brown, NP. 1999. MView: Frequently Asked Questions. Nat Inst for Med Res [Internet]. [cited 2022 Sep]. Available from: http://www.sacs.ucsf.edu/Documentation/seqsoftware/mview/FAQ.html

Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE. 2019. Mouse Genome Database Group. Mouse Genome Database (MGD). Nuc Acids Res [Internet]. [cited 2021 Mar]; 47(D1):D801–D806. Available from: https://doi.org/10.1093/nar/gky1056

Carden S, van der Watt P, Chi A, Ajayi-Smith A, Hadley K, Leaner VD. 2018. A tight balance of Karyopherin β1 expression is required in cervical cancer cells. BMC Cancer [Internet]. [cited 2022 Nov]; 18(1):1123. Available from: https://www.doi.org/10.1186/s12885-018-5044-8

Centers for Disease Control and Prevention. 2020. [Internet]. [cited Nov 2022]. Available from: https://www.cdc.gov/genomics/disease/breast_ovarian_cancer/genes_hboc.htm

Chang W, Dynek JN, Smith S. 2005. NuMA is a major acceptor of poly(ADP-ribosyl)ation by tankyrase 1 in mitosis. Biochem J [Internet]. [cited 2022 Oct]; 391(Pt 2):177-84. Available from: https://doi.org/10.1042/BJ20050885

Citarelli M, Teotia S, Lamb RS. 2010. Evolutionary history of the poly(ADP-ribose) polymerase gene family in eukaryotes. BMC Evol Biol [Internet]. [cited 2022 Nov]; 10:308. Available from: https://doi.org/10.1186/1471-2148-10-308

Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R, et al. 2022. Ensembl. Nuc Acids Res [Internet]. [cited 2022 Jun]; 50(D1): D988–D995. Available from: https://doi.org/10.1093/nar/gkab1049

Daugherty MD, Young JM, Kerns JA, Malik HS. 2014. Rapid Evolution of PARP Genes Suggests a Broad Role for ADP-Ribosylation in Host-Virus Conflicts. PLoS Gen [Internet]. [cited 2021 Jan]; 10(5): e1004403. Available from: https://doi.org/10.1371/journal.pgen.1004403

Di Paola S, Micaroni M, Di Tullio G, Buccione R, Di Girolamo M. 2012. PARP16/ARTD15 is a novel endoplasmic-reticulum-associated mono-ADP-ribosyltransferase that interacts with, and modifies karyopherin-ß1. PLoS One [Internet]. [cited 2022 Oct] ;7(6):e37352. Available from: https://doi.org/10.1371/journal.pone.0037352

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. [Internet]. [cited 2022 Jun]; 39(4):783-791. Available from: https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Gao G, Guo X, Goff SP. 2002. Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein. Science. [Internet]. [cited 2022 Oct]; 297(5587):1703-6. Available from: https://doi.org/10.1126/science.1074276

Gibson BA, Kraus WL. 2012. New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. Nat Rev Mol Cell Biol [Internet]. [cited 2022 Nov]; 13(7):411-24. Available from: https://doi.org/10.1038/nrm3376

Glover JN, Williams RS, Lee MS. 2004. Interactions between BRCT repeats and phosphoproteins: tangled up in two. Trends Biochem Sci [Internet]. [cited 2022 Nov]; 29(11):579-85. Available from: https://doi.org/10.1016/j.tibs.2004.09.010

Gottschalk AJ, Timinszky G, Kong SE, Jin J, Cai Y, Swanson SK, Washburn MP, Florens L, Ladurner AG, Conaway JW, Conaway RC. 2009. Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. Proc Natl Acad Sci U S A [Internet]. [cited 2022 Nov]; 106(33):13770-4. Available from: https://doi.org/10.1073/pnas.0906920106

Guo T, Zuo Y, Qian L, Liu J, Yuan Y, Xu K, Miao Y, Feng Q, Chen X, Jin L, Zhang L, Dong C, Xiong S, Zheng H. 2019. ADP-ribosyltransferase PARP11 modulates the interferon antiviral response by mono-ADP-ribosylating the ubiquitin E3 ligase β-TrCP. Nat Microbiol [Internet]. [cited 2022 Nov]; 4(11):1872-1884. Available from: https://doi.org/10.1038/s41564-019-0428-3

Hajikhezri Z, Darweesh M, Akusjärvi G, Punga T. 2020. Role of CCCH-Type Zinc Finger Proteins in Human Adenovirus Infections. Viruses [Internet]. [cited 2022 Nov] ;12(11):1322. Available from: https://doi.org/10.3390/v12111322

Hakmé A, Wong HK, Dantzer F, Schreiber V. 2008. The expanding field of poly(ADP-ribosyl)ation reactions. 'Protein Modifications: Beyond the Usual Suspects'

Review Series. EMBO Rep. [Internet]. [cited 2022 Nov]; 9(11):1094-100. Available from: https://doi.org/10.1038/embor.2008.191

Han W, Li X, Fu X. 2011. The macro domain protein family: Structure, functions, and their potential therapeutic implications. Mut Res/Rev in Mut Res [Internet]. [cited 2021 Jul]; 727(3):86-103. Available from: https://doi.org/10.1016/j.mrrev.2011.03.001

Hassa PO, Hottiger MO. 2008. The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases. Front Biosci [Internet]. [cited 2022 Nov] ;13:3046-82. Available from: https://doi.org/10.2741/2909

He F, Tsuda K, Takahashi M, Kuwasako K, Terada T, Shirouzu M, Watanabe S, Kigawa T, Kobayashi N, Güntert P, et al. 2012. Structural insight into the interaction of ADP-ribose with the PARP WWE domains. FEBS Lett [Internet]. [cited 2021 May]; 586 (21):3858-3864. Available from: https://doi.org/10.1016/j.febslet.2012.09.009

Huang SM, Mishina YM, Liu S, Cheung A, Stegmeier F, Michaud GA, Charlat O, Wiellette E, Zhang Y, Wiessner S et al. 2009. Tankyrase inhibition stabilizes axin and antagonizes Wnt signaling. Nature [Internet]. [cited 2022 Oct] ; 461(7264):614-20. Available from: https://doi.org/10.1038/nature08356

Hurles M. 2004. Gene duplication: the genomic trade in spare parts. PLoS Biol [Internet]. [cited 2022 Nov]; 2(7):E206. Available from: https://doi.org/10.1371/journal.pbio.0020206

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comp App in Biosci [Internet]. [cited 2022

Jun]; 8(3): 275-282. Available from: https://doi.org/10.1093/bioinformatics/8.3.275

Juszczynski P, Kutok JL, Li C, Mitra J, Aguiar RC, Shipp MA. 2006. BAL1 and BBAP are regulated by a gamma interferon-responsive bidirectional promoter and are overexpressed in diffuse large B-cell lymphomas with a prominent inflammatory infiltrate. Mol Cell Biol [Internet]. [cited 2022 Oct]; 26(14):5348-59. Available from: https://doi.org/10.1128/MCB.02351-05

Jwa M, Chang P. 2012. PARP16 is a tail-anchored endoplasmic reticulum protein required for the PERK- and IRE1α-mediated unfolded protein response. Nat Cell Biol [Internet]. [cited 2022 Oct]; 14(11):1223-30. Available from: https://doi.org/10.1038/ncb2593

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Gen Res [Internet]. [cited 2021 Mar]; 12(6):996-1006. Available from: https://genome.ucsc.edu/cgi-bin/hgGateway

Kolishovski G, Lamoureux A, Hale P, Richardson JE, Recla JM, Adesanya O, Simons A, Kunde-Ramamoorthy G, Bult CJ, 2019. The JAX Synteny Browser for mouse-human comparative genomics. Mamm Genome 30 [Internet]. [cited 2022 Oct]; 353–361. Available from: Available from: https://doi.org/10.1007/s00335-019-09821-4

Kraus WL. 2008. Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. Curr Opin Cell Biol [Internet]. [cited 2022 Oct]; 20(3):294-302. Available from: https://doi.org/10.1016/j.ceb.2008.03.006

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol [Internet]. [cited 2022 Oct]; 35(6):1547-1549. Available from: https://doi.org/10.1093/molbev/msy096 and https://www.megasoftware.net/

Kwon WK, Jang HJ, Lee JE, Park YH, Ryu JM, Yu J, Jang JH, Kim JW. 2022. Discovery of BRCA1/BRCA2 founder variants by haplotype analysis. Cancer Genet [Internet]. [cited 2022 Nov]; 266-267:19-27. Available from: https://doi.org/10.1016/j.cancergen.2022.05.042

Letunic I, Khedkar S, Bork P. 2021. SMART: recent updates, new developments and status in 2020. Nuc Acids Res [Internet]. [cited 2021 Jul]; 49(D1):D458–D460. Available from: https://doi.org/10.1093/nar/gkaa937

Levaot N, Voytyuk O, Dimitriou I, Sircoulomb F, Chandrakumar A, Deckert M, Krzyzanowski PM, Scotter A, Gu S, Janmohamed S, et al. 2011. Loss of Tankyrase-mediated destruction of 3BP2 is the underlying pathogenic mechanism of cherubism. Cell [Internet]. [cited 2022 Nov]; 147(6):1324-39. Available from: https://doi.org/10.1016/j.cell.2011.10.045

Li B, Navarro S, Kasahara N, Comai L. 2004. Identification and biochemical characterization of a Werner's syndrome protein complex with Ku70/80 and

poly(ADP-ribose) polymerase-1. J Biol Chem [Internet]. [cited 2022 Nov]; 279(14):13659-67. Available from: https://doi.org/10.1074/jbc.M311606200

Li N, Chen J. 2014. ADP-Ribosylation: Activation, Recognition, and Removal. Mol Cell [Internet]. [cited 2021 Feb]; 37(1):9-16. Available from: http://dx.doi.org/10.14348/molcells.2014.2245

Li X, Erden O, Li L, Ye Q, Wilson A, Du W. 2014. Binding to WGR Domain by Salidroside Activates PARP1 and Protects Hematopoietic Stem Cells from Oxidative Stress. Anti-Red Sig [Internet]. [cited 2021 May]; 20(12):1853-1865. Available from: https://doi.org/10.1089/ars.2013.5600

Loraine AE, Blakley IC, Jagadeesan S, Harper J, Miller G, Firon N. 2015. Analysis and visualization of RNA-seq expression data using RStudio, Bioconductor, and Integrated Genome Browser. Met Mol Biol [Internet]. [cited 2021 Mar]; 1284:481-501. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4387895/. Link to supplemental data including R code: https://bitbucket.org/lorainelab/tomatopollen/src/master/ RNA-Seq MDS Plot and Cluster Dendrogram Plot were obtained from use of RStudio Version 4.0.5 from March 2021

Lord CJ, Ashworth A. 2008. Targeted therapy for cancer using PARP inhibitors. Curr Opin Pharmacol [Internet]. [cited 2022 Oct]; 8(4):363-9. Available from: https://doi.org/10.1016/j.coph.2008.06.016

Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. Science. [Internet]. [cited 2022 Nov]; 290(5494):1151-5. Available from: https://doi.org/10.1126/science.290.5494.1151

MacPherson L, Tamblyn L, Rajendra S, Bralha F, McPherson JP, Matthews J. 2013. 2,3,7,8-Tetrachlorodibenzo-p-dioxin poly(ADP-ribose) polymerase (TiPARP, ARTD14) is a mono-ADP-ribosyltransferase and repressor of aryl hydrocarbon receptor transactivation. Nucleic Acids Res [Internet]. [cited 2022 Oct]; 41(3):1604-21. Available from: https://doi.org/10.1093/nar/gks1337.

Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. 2013. Gene duplication as a major force in evolution. J Gen [Internet]. [cited 2021 Jan]; 92(1):155-161. Available from: https://www.ias.ac.in/article/fulltext/jgen/092/01/0155-0161

Mao R, Nie H, Cai D, Zhang J, Liu H, Yan R, Cuconati A, Block TM, Guo JT, Guo H. 2013. Inhibition of hepatitis B virus replication by the host zinc finger antiviral protein. PLoS Pathog [Internet]. [cited 2022 Nov]; 9(7):e1003494. Available from: https://doi.org/10.1371/journal.ppat.1003494

Maynard KB, Smith SA, Davis AC, Trivette A, Seipelt-Thiemann RL. 2014. Evolutionary analysis of the mammalian M1 aminopeptidases reveals conserved exon structure and gene death. Gene [Internet]. [cited 2021 Jan]; 552(1):126-132. Available from: https://doi.org/10.1016/j.gene.2014.09.025

Mehrotra P, Riley JP, Patel R, Li F, Voss L, Goenka S. 2011. PARP-14 functions as a transcriptional switch for Stat6-dependent gene activation. J Biol Chem [Internet]. [cited 2022 Oct]; 286(3):1767-76. Available from: https://doi.org/10.1074/jbc.M110.157768

Müller S, Möller P, Bick MJ, Wurr S, Becker S, Günther S, Kümmerer BM. 2007. Inhibition of filovirus replication by the zinc finger antiviral protein. J Virol. [Internet]. [cited 2022 Nov]; 81(5):2391-400. Available from: https://doi.org/10.1128/JVI.01601-06

Nguewa PA, Fuertes MA, Valladares B, Alonso C, Pérez JM. Poly(ADP-ribose) polymerases: homology, structural domains and functions. 2005. Novel therapeutical applications. Prog Biophys Mol Biol [Internet]. [cited 2022 Oct]; 88(1):143-72. Available from: https://doi.org/10.1016/j.pbiomolbio.2004.01.001

Obayashi T, Kagaya Y, Aoki Y, Tadaka S, Kinoshita K. 2019. COXPRESdb v7: a gene coexpression database for 11 animal species supported by 23 coexpression platforms for technical evaluation and evolutionary inference. Nuc Acids Res [Internet]. [cited 2021 Mar]; 47(D1):D55-D62. Available from: https://doi.org/10.1093/nar/gky1155

Ocampo Daza D, Sundström G, Bergqvist CA, Duan C, Larhammar D. 2011. Evolution of the insulin-like growth factor binding protein (IGFBP) family. Endocrin [Internet]. [cited 2022 Oct]; 152(6):2278-89. Available from: https://doi.org/10.1210/en.2011-0047

Ohno S. 1970. Evolution by gene duplication. New York, USA. Springer-Verlag. 160 pp.

Papadopoulos JS, Agarwala R. 2007. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics [Internet]. [cited 2022 Nov]; 23(9):1073-9. Available from: https://doi.org/10.1093/bioinformatics/btm076

Parsons JL, Dianova II, Allinson SL, Dianov GL. 2005. Poly(ADP-ribose) polymerase-1 protects excessive DNA strand breaks from deterioration during repair in human cell extracts. FEBS J. [Internet]. [cited 2022 Oct]; 272(8):2012-21. Available from: https://doi.org/10.1111/j.1742-4658.2005.04628.x

Perina D, Mikoč A, Ahel J, Ćetković H, Žaja R, Ahel I. 2014. Distribution of protein poly(ADP-ribosyl)ation systems across all domains of life. DNA Repair (Amst) [Internet]. [cited 2022 Dec]; 23:4-16. Available from: https://doi.org/10.1016/j.dnarep.2014.05.003

Raval-Fernandes S, Kickhoefer VA, Kitchen C, Rome LH. 2005. Increased susceptibility of vault poly(ADP-ribose) polymerase-deficient mice to carcinogen-induced tumorigenesis. Cancer Res [Internet]. [cited 2022 Oct]; 65(19):8846-52. Available from: https://doi.org/10.1158/0008-5472.CAN-05-0770

Riffell JL, Lord CJ, Ashworth A. 2012. Tankyrase-targeted therapeutics: expanding opportunities in the PARP family. Nat Rev Drug Discov [Internet]. [cited 2022 Nov]; 11(12):923-36. Available from: https://doi.org/10.1038/nrd3868

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. [Internet]. [cited 2022 Nov]; 26(1):139-40. Available from: https://doi.org/10.1093/bioinformatics/btp616

Ruscetti T, Lehnert BE, Halbrook J, Le Trong H, Hoekstra MF, Chen DJ, Peterson SR. 1998. Stimulation of the DNA-dependent protein kinase by poly(ADP-ribose) polymerase. J Biol Chem [Internet]. [cited 2022 Oct]; 273(23):14461-7. Available from: https://doi.org/10.1074/jbc.273.23.14461

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Bio Evol [Internet]. [cited 2022 Jun]; 4(4):406-425. Available from: https://doi.org/10.1093/oxfordjournals.molbev.a040454

Schreiber V, Dantzer F, Ame JC, de Murcia G. 2006. Poly(ADP-ribose): novel functions for an old molecule. Nat Rev Mol Cell Biol [Internet]. [cited 2022 Nov]; 7(7):517-28. Available from: https://doi.org/10.1038/nrm1963

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol [Internet]. [cited 2021 Mar]; 7:539. Available from: https://doi.org/10.1038/msb.2011.75

Smith JR, Hayman GT, Wang SJ, Laulederkind SJF, Hoffman MJ, Kaldunski ML, Tutaj M, Thota J, Nalabolu HS, Ellanki SLR, et al. 2019. The Year of the Rat: The Rat Genome Database at 20: a multi-species knowledgebase and analysis platform. Nuc Acids Res [Internet]. [cited 2022 Jun]; 48(D1):D731-D742. Available from: https://doi.org/10.1093/nar/gkz1041

Smith S, de Lange T. Tankyrase promotes telomere elongation in human cells. 2000. Curr Biol [Internet]. [cited 2022 Oct]; 10(20):1299-302. Available from: https://doi.org/ 10.1016/s0960-9822(00)00752-1

Smith S, Giriat I, Schmitt A, de Lange T. 1998. Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. Science [Internet]. [cited 2022 Nov] ;282(5393):1484-7. Available from: https://doi.org/10.1126/science.282.5393.1484

Stecher G, Tamura K, Kumar S. 2020. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Mol Bio Evol [Internet]. [cited 2022 Jun]; 37(4):1237-1239. Available from: https://doi.org/10.1093/molbev/msz312 and https://www.megasoftware.net/

Stothard P, 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques [Internet]. [cited 2022 Oct]; 28(6):1102-1104. Available from: https://doi.org/10.2144/00286ir01

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol [Internet]. [cited 2022 Jun]; 38(7):3022-3027 Available from: https://doi.org/10.1093/molbev/msab120 and https://www.megasoftware.net/

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses mentioned in this project were obtained from: [RNA-Seq Data-Median gene-level TPM by tissue] the GTEx Portal on 03/01/21 and/or dbGaP accession number phs000424.v8.p2 on 03/01/2021 [Internet]. [cited 2021 Mar]. Available from: https://www.gtexportal.org/home/datasets

Data was also obtained from: [GTEx Multi Gene Query Database]. [cited 2022 Jun]. Accessed on 06/20/22. Available from: https://www.gtexportal.org/home/multiGeneQueryPage

Timinszky G, Till S, Hassa PO, Hothorn M, Kustatscher G, Nijmeijer B, Colombelli J, Altmeyer M, Stelzer EH, Scheffzek K, Hottiger MO, Ladurner AG. 2009. A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. Nat Struct Mol Biol [Internet]. [cited 2022 Nov]; 16(9):923-9. Available from: https://doi.org/0.1038/nsmb.1664

Tostivint H, Ocampo Daza D, Bergqvist CA, Quan FB, Bougerol M, Lihrmann I, Larhammar D. 2014. Molecular evolution of GPCRs: Somatostatin/urotensin II receptors. J Mol Endocrin [Internet]. [cited 2022 Oct]; 52(3):T61-86. Available from: https://doi.org/10.1530/JME-13-0274

Tuncel H, Tanaka S, Oka S, Nakai S, Fukutomi R, Okamoto M, Ota T, Kaneko H, Tatsuka M, Shimamoto F. 2012. PARP6, a mono(ADP-ribosyl) transferase and a negative regulator of cell proliferation, is involved in colorectal cancer development. Int J Oncol [Internet]. [cited 2022 Oct]; 41(6):2079-86. Available from: https://doi.org/10.3892/ijo.2012.1652

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. 2015. Proteomics. Tissue-based map of the human proteome. Science [Internet]. [cited 2022 Dec]; 347(6220):1260419. Available from: https://doi.org/10.1126/science.1260419

Vyas S, Chesarone-Cataldo M, Todorova T, Huang YH, Chang P. 2013. A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. Nat Com [Internet]. [cited 2021 Feb]; 4:2240. Available from: https://doi.org/10.1038/ncomms3240

World Health Organization. 2020. [Internet]. [cited 2022 Oct]. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer

Yan Q, Xu R, Zhu L, Cheng X, Wang Z, Manis J, Shipp MA. 2013. BAL1 and its partner E3 ligase, BBAP, link Poly(ADP-ribose) activation, ubiquitylation, and double-strand DNA repair independent of ATM, MDC1, and RNF8. Mol Cell Biol [Internet]. [cited 2022 Oct]; 33(4):845-57. Available from: https://doi.org/10.1128/MCB.00990-12

Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, Bhai J, Campbell LI, Carbajo Martinez M et al. 2022. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. Nuc Acids Res [Internet]. [cited 2022 Oct]; 50(D1):D996-D1003. Available from: https://doi.org/10.1093/nar/gkab1007

Yu M, Schreek S, Cerni C, Schamberger C, Lesniewicz K, Poreba E, Vervoorts J, Walsemann G, Grötzinger J, Kremmer E, et al. 2005. PARP-10, a novel Myc-interacting protein with poly(ADP-ribose) polymerase activity, inhibits transformation. Oncogene [Internet]. [cited 2022 Nov]; 24(12):1982-93. Available from: https://doi.org/10.1038/sj.onc.1208410

Yu X, Chini CC, He M, Mer G, Chen J. 2003. The BRCT domain is a phospho-protein binding domain. Science [Internet]. [cited 2022 Nov]; 302(5645):639-42. Available from: https://doi.org/10.1126/science.1088753

Zhang J. 2003. Evolution by gene duplication: an update. Trends Ecol Evol [Internet]. [cited 2022 Nov]; 18(6), 292–298. Available from: https://content.csbs.utah.edu/~rogers/tch/bio5410/Readings/Zhang-TRE-18-292.pdf

Zuckerkandl E, Pauling L. 1965. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by Bryson V and Vogel HJ. Acad Pres [Internet]. [cited 2022 Jun]; 97-166. Available from: https://doi.org/10.1016/B978-1-4832-2734-4.50017-6